DOCUMENT RESUME

ED 145 956 PS 009 635

AUTHOR Murray, Harry William
TITLE Early Intervention in the Context of Family Characteristics.
SPONS AGENCY Education Commission of the States, Denver, Colo.; Office of Child Development (DHEW), Washington, D.C.
PUB DATE Apr 77
NOTE 74p.; Tables and figures have been filmed from best available copy; Paper presented at the Annual Meeting of the American Orthopsychiatric Association (New York, New York, April 12-16, 1977)

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS Birth Order; Correlation; *Early Childhood Education; Early Experience; *Family Characteristics; Family Structure; *Intelligence Quotient; *Intervention; Longitudinal Studies; *Low Income Groups; Parental Background; Preschool Programs; Socioeconomic Status; *Statistical Analysis

ABSTRACT

This paper compares results from 11 longitudinal studies to determine whether preschool education improves the IQ of low income children when their family situations are taken into account and if so, how long the effects last. Taken together, the 11 studies represent a sample of 1,645 children -- of whom 61% are male, 87% black, and 75% former participants in experimental preschools. Many analyses of this paper include only portions of the total sample, since not all projects collected information pertinent to every analysis. The dependent variable in all analyses is Stanford-Binet IQ; the family structure variables investigated are socioeconomic status (SES), mother's education, family size, and birth order. The results indicate that there is a significant probability that lower class children who attend a preschool will have higher IQ's than those who do not, for at least 3 years after preschool. Findings also show that for lower class populations, significant correlations exist between mother's education, SES, family size, and birth order and IQ, but that the combined effect of these variables on the percent of variance explained for IQ is slight. Finally, little evidence is found that preschools affect the correlations between family variables and IQ. (Author/JMB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED145956

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Early Intervention in the Context of Family Characteristics

Harry William Murray

1/30/77

paper for presentation at the

American Orthopsychiatric Association Annual Meeting

April 16, 1977

BEST COPY AVAILABLE

PS 009635

The debate over the effectiveness of preschool educational programs for low income children has wandered in and out of the national spotlight since the 1969 Westinghouse Report concluded that Head Start had no long term effects on cognitive functioning. The major national evaluations of compensatory education have found disappointingly few clear results upholding the effectiveness of such educational programs in increasing IQ and academic performance. Arguments against the effectiveness of preschools have been widely publicized: Jensen argues that compensatory education has failed to raise IQ on a long term basis because IQ is largely determined by genetics and, hence, is basically unalterable;[1] Bronfenbrenner argues (more convincingly) that most preschools fail to raise IQ or achievement in the long run because they affect only a small part of the child's total environment, and that any serious effort to raise IQ would entail a more massive attempt to improve the environment of low income children.[2] The debate is still going on, but the "current wisdom" is that preschool programs for low income children boost IQ for a year or two but subsequently do not prevent it from falling back to around the original level.

Among a number of smaller studies of experimental preschool programs for low income children, which have collected longitudinal data on the children's development, some support the current wisdom, while others indicate the possibility that preschool programs can cause IQ gains which remain significant for a number of years. Only a few of these studies have been mentioned in the debate over Head Start, and even these have often been dismissed with remarks such as: "Only one or two small studies claim appreciable differences beyond first grade."[3] Most of these studies have been reviewed several times, but their results have never been systematically compared in anything approaching a rigorous manner. This paper attempts a systematic comparison of results using

---

This paper was prepared under a grant from the Office of Child Development through the Education Commission of the States under the direction of Dr. Irving Lazar.

the data from eleven such studies. Taken together, these studies represent a sample of 1,645 children-of whom 61% are male, 87% Black, and 75% former attendants of experimental preschools. Many analyses in this paper only include portions of the total sample, since not all projects collected the information for every analysis.

This paper addresses the question of whether preschool improves the IQ of low income children when their family situations are taken into account, and if it does, for how long does this effect last? The dependent variable is IQ as measured by the Stanford-Binet scale, and the family structure variables which will be investigated are socioeconomic status (SES), mother's education, family size and birth order.

Family variables are integrally tied up with the question of the effectiveness of preschool education simply because numerous theoretical and empirical studies have shown that a wide range of family variables (including socioeconomic status, family size and structure, family process, and parental attitudes)[5] can affect both IQ and academic achievement - which are, of course, the primary criteria by which preschool effectiveness is generally measured. There are two basic (although not total distinct) ways in which family variables can affect the results of evaluations of early childhood education. First, the family variables which affect IQ are not always properly accounted for in the design of the evaluation and hence, may bias results one way or another. For example, if the sample selection procedures result in significantly different types of families being assigned to the experimental group as opposed to the control or comparison group,* this difference, rather than any effects of the preschool program, could be the true cause of the results of the evaluation.[6] A second way in which family variables can affect the measured results of preschool programs is through

---

* For simplicity, the term "control group" will be used in this article to mean "control or comparison group."

their interaction with the programs: different types of children could be affected in different ways by different types of programs. Herzog, for example, finds that "the less they have, the less they learn," i.e., that her program was more effective with the less deprived children.[7] Other studies of Head Start programs have found little interaction between family variables and Head Start effects.[8]

Since this paper deals primarily with the effects of attending a preschool versus not attending one (rather than with a comparison of different types of preschools), it emphasizes the former way, in which family variables affect evaluations, rather than the latter in which family variables interact with different types of programs. However, since the two questions are interrelated, some light will also be shed on the latter point.

This analysis is part of a larger longitudinal research effort funded by the Office of Child Development through the Education Commission of the States. The project is conducted by a consortium of twelve principal investigators who have conducted longitudinal research on early childhood education projects for low income people. The consortium is chaired by Dr. Irving Lazar and the members are: Kuno Beller, Martin Deutsch, Cynthia Deutsch, Ira Gordon, Susan Gray, Merle Karnes, Phyllis Levenstein, Louise Miller, Francis Palmer, David Weikart, Myron Woolman and Edward Zigler. This group is currently engaged in two interrelated efforts to provide better information on the actual effects of preschools on low income children:

(1) pooling the original longitudinal data collected by the individual projects (from which the current analysis is drawn, and (2) collecting current year followup data from all projects in identical format for use in both pooled and separate analyses. Since this paper represents one of the first products of the consortium and is an analysis of only one dependent variable,

the methodology for integrating the results of the projects is relatively conservative and emphasizes respect for the integrity of each project.

METHODOLOGY

A. General

Raw test and demographic data were duplicated at each project and transmitted to the central office. This raw data was then checked for accuracy and, where necessary, corrected after consultation with personnel at the project from which it came. The IQ test and demographic data were then transformed to a common format to facilitate analysis.

For purposes of this analysis a case was defined as any subject whose test score was transmitted to the central office. In some instances, this definition varies slightly from the definition of a case used by the original project. Hence, the results of analyses as reported here may not always coincide exactly with those reported previously by the original projects.

Of the data transmitted by all consortium members eleven data sets were selected for analysis in this paper:

- Ira Gordon's Parent Education/Back Yard Center Project, which provided two types of home-visitor, parent-focused intervention during the middle '60's to zero-to-three year old migrant children in the Gainesville, Florida area.
- Susan Gray's Family Oriented Home Visitor Project (FOHV), which provided home visits to one and two year old children in Nashville in the early '70's.
- Merle Karnes' longitudinal project which provided and compared five types of center-based education (Bereiter-Englemann, Montessori, traditional, community-integrated, and the Karnes' GOAL approach) to four year old children in Champaign-Urbana, Illinois in the middle '60's.
- Merle Karnes' pre-post groups, which actually comprise a number of separate studies of different types of intervention for three and four year olds in the same environment as the longitudinal project. (Since there was no longitudinal

follow-up on these groups, their data is used only for estimating the correlation of the family variables with IQ).

- The first five waves of Phyllis Levenstein's Verbal Interaction Project, which provided home visitors to mother-child dyads beginning at age two on Long Island in the late 1960's and early 1970's.
- Louise Miller's Experimental Variation of Head Start Curricula, which provided and compared four types of programs for four year olds (Bereiter-Englemann, Montessori, DARCEE, and traditional) in the middle sixties in Louiseville, Kentucky.
- Francis Palmer's Harlem Training Project, which provided one to one center based intervention for two and three year old Black boys in Harlem in the middle sixties.
- David Weikart's Perry Preschool Project, which provided a combination of center and home-based intervention to very low IQ (less than 85) three and four year old children in Ypsilanti, Michigan in the early sixties.
- David Weikart,s Curriculum Demonstration Project (CD) which provided and compared three approaches to intervention for very low IQ three year olds (Bereiter-Englemann, traditional, and Weikart's Cognitively Oriented Curriculum which had benn used in the Perry project) in the middle sixties, again in Ypsilanti.
- David Weikart's Carnegie Infant Program, which provided home visitors to zero-to-two year olds in the late sixties in Ypsilanti.
- Susan Gray's Early Training Project (ETP), which gave both center- and home-based intervention to preschool Black children in Murfreesboro, Tenn. in the early 1960's.

Four projects which are involved in the consortium's effort are excluded from analysis at this time for the following reasons:

- Kuno Beller's project, which provided a center-based program for four year olds in Philadelphia in the early sixties and which showed significant results on the Stanford-Binet IQ test through fourth grade, was excluded due to lack of adequate demographic data.

- Martin and Cynthia Deutsch's Institute for Developmental Studies project, which provided schooling for Harlem youngsters from age four through the third grade, was excluded because the data were received too late for adequate validation and preparation.
- Myron Woolman's Micro-Social Learning System, which provided a unique systems-oriented approach to education for four year old migrant children in Vineland, New Jersey, was excluded because it did not use the Binet as a measure.
- Edward Zigler's study of the effects of Head Start and Followthrough on children in New Haven, which has found significant results as late as eighth grade, was excluded because it did not use the Binet as a measure.

Our inability to use the data from these projects is quite unfortunate, particularly since two of them have shown evidence of very long term effects. These projects will be included in future analyses.

It should be noted that, since it is a secondary analysis, this paper is addressed toward questions which the projects from which it draws information were not specifically designed to answer. Several of the projects (Karnes longitudinal, Miller, and Curriculum Demonstration), in fact, were aimed specifically at comparing different approaches. All of the other projects except the Perry Preschool and the Carnegie Infant program compared variations of programs (either in terms of age at which the program was administered or in terms of variations of the curriculum or approach or in terms of both). However, this analysis deals primarily with the effect of participating in a preschool program in the context of family variables and does not make a systematic attempt to distinguish among types of preschool approaches. Therefore, these variations are classified together as "having received experimental preschool education" and are contrasted only against controls who had not received such an education rather than against each other. Thus, it must be remembered that the emphases in this paper are somewhat

different from the emphases in the minds of the original researchers.

B. The Stanford-Binet IQ Test

A specific example of the difference in emphasis between this paper and the original investigations is the selection of the Stanford-Binet IQ Scale as the dependent variable for analysis. The Binet was chosen because it was the test most commonly used by the consortium projects and because it has had a fairly large scale standardization. However, the qualities tested by the Binet are in some cases only peripherally related to what the particular preschool program in a project attempted to accomplish. In all projects the Binet was only one of many tests used to evaluate the program. The Binet evaluates a wide range of abilities, only a few of which may have been the targets of a given project. Thus, the project could achieve a significant improvement in the abilities toward which it was aimed, but this improvement may not have been enough to provide a significant change in the Binet score because the Binet also tapped other abilities which the project had not attempted to improve. A perusal of individual project reports reveals that in many instances the researchers found other tests to be far more relevant than IQ to the purposes of the project. In fact these tests often yielded more favorable results than the Binet. (Unfortunately, however, since such tests were more specific to the goals of a project, none of these tests were used by many projects, and, hence, they cannot form a solid basis for comparison of the projects.)

In using the Binet as a dependent variable, this paper is employing a very stringent criterion of success for these projects, and therefore a failure to improve the Binet IQ cannot be taken as a failure of a program to accomplish its goals. On the other hand, however, the success of a project in improving IQ means that it has accomplished one goal which our society deems important.

This paper is also adopting a conservative tactic by comparing only Binet IQ's (rather than comparing results on Binet, WISC, Cattell and other IQ tests) due to the problems inherent in comparing one IQ test with another. This approach results in the loss of early scores in the Gordon, FOHV, VIP, and Carnegie projects, as well as several followup testing periods in the VIP project. Problems of comparability across projects still occur because of differences in testing situations (e.g., differences due to sex and race of tester, or extent of efforts to let the child "adjust" to the testing situation). However, such differences are assumed to be negligible unless otherwise noted in the text.

IQ scores are reported on the basis of the 1960 deviation IQ tables because these were the scores reported by the individual projects. Caution should be used in interpreting the IQ levels and in comparing IQ across ages because the 1972 restandardization of the Binet indicated that the mean Binet IQ's from the 1960 tables are not 100 at every age, but vary by age, from 110.7 at age three years to 101.9 at age ten.[9] If the results of the 1972 restandardization are in fact more accurate than the results of the previous standardization, the Binet IQ's according to the 1960 tables overstate the IQ's for the younger children by over 10 points; therefore, a child's IQ could naturally be expected to decline steadily between the ages of three and ten solely as an artifact of the earlier standardization. This problem does not, of course, affect correlations.

For the purpose of comparing testing periods across projects, one testing period "slot" was created for every year of age (e.g., the first slot ranged from eighteen to twenty-nine months - the midpoint being twenty-four months, etc.)* Testing periods were then assigned to the slots in which their mean

* Thus, all children in one project's testing period were always placed in the same slot. This procedure might seem too obvious to mention; however, specification of the procedure was necessary because some projects defined testing

chronological age fell. For example, any testing period with a mean chronological age between thirty and forty-one months would be compared with any other projects' testing periods where the mean chronological age was between thirty and forty-one months. This procedure ensures that only tests which were given at approximately the same age are compared with each other. Even this procedure, however, may not be enough to completely eliminate bias due to children being tested at different ages since, as reported by Kamin, such bias can occur even for children who differ by only a few months in age:

> The chronological ages are given in months, and the 53 6-year-olds ranged from 6 years, 0 months to 6 years, 10 months. Within this very restricted range of ages, the correlation between age and IQ was a significant -.35...This stable and significant correlation was clearly an artifact built into the Stanford-Binet test. 10

However, the use of testing period slots to group children of similar chronological age (within one year of each other) minimizes the age bias to the greatest extent possible.**

C. Family Variables

The analysis of family variables is another important element of this paper which was peripheral to most of the projects. About half of the projects collected demographic information mainly to describe their sample or as part of the selection and intake process, rather than for the purposes of analysis. Thus, for a number of projects, the demographic information had never even been put into machine-readable form until the advent of this joint endeavor.

---

period by chronological age regardless of time in the school year, while others defined testing period by time of the school year (e.g., post first grade).

** An alternative procedure would have been to define testing period slots by number of years after the termination of the preschool; however, this would have entailed comparing IQ across different ages and, hence, was rejected.

Comparing demographic or family variables across projects is, if anything, more difficult than comparing test scores.

A major problem arises from the difference in definitions of the variables. For instance, among the projects the definition of siblings (and, hence, of birth order) ranged from "any other child living in the home" to "any other child of the mother." Fortunately, in this instance, the Harlem Training Project and the Early Training Project had collected both definitions and comparison revealed no significant differences between the definitions.* Therefor, it was decided that no adjustment was necessary for the difference in definition of this variable.

Even for variables where there were no definitional problems, such as parental education (measured in grades completed) or occupation (generally recorded in a form which could be or already was coded into the Hollingshead occupation categories), there arises the question of whether the same value on the same variable means the same thing for people in different environments. Did a twelfth grade education for a Black mother in Murfreesboro, Tenn. in the early sixties mean the same thing as a twelfth grade education for a mother in Freeport, Long Island in the early seventies? Since this paper deals primarily with parallel analyses (i.e., looking at relationships within each project and then comparing the relationships across projects) this question need not be rigorously addressed at this time.

A final complication in comparing demographic variables across projects is the fact that many demographic variables change over time (e.g., SES, father presence, number of siblings). Therefore, if demographic data is to be compared

---

* Specifically, it was found that: (a) the different definitions only caused a change in the value of the variable in less than 5% of the cases, and (b) there was no significant difference in means between the definitions (t-test), and (c) the different definitions were highly correlated ($r > .9$).

to IQ at a number of points in time, the demographic data should be recollected at each time point, as well as the IQ data, if the research is to be totally accurate.* Often, however, the demographic data is collected at only one time point and is assumed to be relatively stable. The particular problem which the variability of demographic variables over time raises for comparing different studies is that the studies collected the demographic information at slightly different ages, and this difference could cause some biases (e.g., a newborn child would be far less likely to have younger siblings than a child of four; thus, data collected for newborn chiddren would be likely to yield a smaller average number of siblings than data collected for four year olds). This possible source of bias was pretty well controlled in this analysis since the demographic information for most of the children was collected between the ages of two and four.

A related problem arises from collecting the data for number of siblings at an early age - this variable may not represent "completed family size" since siblings may have been born after the data were collected. Hence, the variable may not be as strongly correlated with IQ as "completed family zize" would be.

The demographic variables thus presented a number of problems for an analysis which would compare them across projects, and the existence of these problems was a major factor in deciding to emphasize a detailed analysis of parallel comparisons.

The family variables to be used in this analysis were selected on the basis of two criteria: (1) their theoretical importance and the prior empirical demonstration of their relation to IQ and (2) whether they were collected by most of the projects in reasonably comparable form. The second criterion

* More generally, the question arises as to when an IQ test should be given to compare to demographic variables collected at a certain age. Should they be concurrent or is there a time-lag phenomenon?

was decisive. Three variables which may have important relationships to IQ - family income, source of income, and father presence - are excluded from the analysis because of this criterion*: Race was not analyzed because an overwhelming majority (87%) of the sample were Black.

Four variables are considered in this study: head of household socioeconomic status (SES), mother's education, number of siblings, and birth order. These variables were collected by most of the eleven projects. The following paragraphs discuss their precise definition for this study and cite some prior theoretical and empirical demonstrations of their relation to IQ.

SES is defined for this study as the Hollingshead Two Factor Index of Social Position for the head of the household.** SES is perhaps the family variable most commonly recognized as being associated with IQ and achievement.

> One set of background factors that is shown by nearly every study of background to have relatively high association with measures of school performance is socioeconomic status. This generalization holds over a wide variety of outcome measures and over a wide variety of ways in which SES is measured. [11]

Project TALENT reports high multiple correlations between a number of achievement measures and twenty-five items representing social class. The multiple correlations (R's) ranged from .24 to .69, depending on the type of test.[12]

---

* Income was rejected because only three projects had collected data and these had collected it in categories which were incompatible across projects. Moreover, socioeconomic status is a reasonable proxy for income. Welfare status and father presence were both collected by a majority of the projects but were rejected for this analysis for three reasons: (1) differences in the precise definitions used by different projects, (2) the high variability of these variables over time, and (3) verbal reports from the investigators who had collected these data because the interviewee's response may be affected by the social stigma attached to the variables.

** This scale was selected because ten of the projects had collected the necessary educational and occupational data while few projects had collected the income data which would have been necessary for a more complex three factor index.

Whiteman and Deutsch report a multiple correlation of .49 between reading scores and six measures of cultural deprivation.[13] Milner finds a simple Pearson correlation of .86 between a composite score from a number of reading readiness tests and a social class index.[14] Although the latter study dealt only with Blacks, all of the above studies dealt with a wide range of social classes, in contrast to the present effort. Jencks concludes:

> Taking all the evidence together, however, we estimate that a family's economic status probably correlates about 0.35 with the childrens' test scores. [15]

Mother education for this analysis is defined simply as the number of grades completed. Like SES, this variable has been shown to be positively correlated with achievement and IQ for large national samples. Bachman reports a correlation of .25 between a scale of mother's education and reading achievement. Husen reports a similar correlation of .28 between math scores for thirteen year olds and number of years of mother's education. Project TALENT finds a correlation of about .27 with the information component of their ability tests.[16] Thus, mother education seems to be strongly related to IQ and achievement measures, although not as strongly as more global (or multiple) measures of SES.

Number of siblings and birth order are dealt with together since they are often treated as a unit in both theoretical and empirical reviews of their relation to IQ. Both are treated as continuous variables in this analysis; definitional problems were discussed above. Number of siblings was selected as the functional expression for family size over two alternate functional forms for several technical reasons.* The empirical findings on the relation-

---

* The two alternative forms of the variable were total family size and the inverse of the number of children. Family size was rejected for two reasons: (1) There were great differences in definition which would hinder comparison across projects (e.g., "number of persons living in the household" vs. "number in the family." (2) Family size as a variable does not differentiate between the effect of an additional adult as opposed to the effect of an

ship of family size and birth order to IQ are well summarized by Terhune:

> Our review of the literature has shown that the negative correlation between intelligence and family size is a reliable one. Only in special groups does the relation fail to appear, viz. in higher socioeconomic groups and among those raised on farms. The correlation is, however, small, usually ranging from -.20 to -.30. Controlling for social class generally reduces the correlation to -.12 to -.16 or further. Consequently, only about one to three percent of variance is associated with family size differentials, and of that limited variance, part is due to heredity (duller people having larger families) and part is due to environment...Part of the family size effect is due to birth order, for large families yield children of higher birth orders, who tend to be intellectually disadvantaged. The explanation for that is probably that later-borns generally received less parental attention...Beyond birth order, there seem to be effects of sheer number of children in the family, from which even first-borns suffer. [17]

D. Overview of the Analysis

As the first step in the analysis, three possible sources of bias are investigated: (1) basic differences among children in different projects, (2) the effects of sample selection and procedures for assignment to experimental or control groups in each project, and (3) the effects of attrition in each project.

The first possible source of bias can, perhaps, best be phrased as a question: Although all projects claimed to serve "disadvantaged" children, to what extent did "disadvantaged" mean the same thing across projects? This question is analyzed by comparing the mean scores of each project's children on measures of SES, maternal education and pre-test Binet IQ using one way analysis of variance to test for overall differences and the Scheffe test to determine which projects are significantly different from each other.

The analyses of sample selection and attrition are of interest primarily to methodologists; therefore, only the results which relate to the reader's evaluation of our analyses are presented in the body of the article, while the analyses themselves are presented in the Appendix.

---

additional child despite the fact that most of the theoretical literature either makes such a distinction (e.g., Zajonc and Markus) or deals with only the effect of additional children (e.g., Cicirelli, and Marjoribank and Walberg). The inverse of the number of children (recommended by Marjoribank and Walberg on the rationale that what effects IQ is the amount of parental attention, which is a function of the inverse of the number of children) was rejected on the basis of slightly lower correlations with IQ than the simpler formulation (# of siblings).

The analysis of the effects of preschool education in the context of family variables is presented in three stages.

In the first stage, the IQ differences between experimental and control groups will be analyzed at each age for each project by means of t-tests. The probabilities that the experimentals had significantly higher IQ scores than the controls at a specific number of years after leaving the program are grouped by length of time after the termination of the preschool. Using the technique recommended by Darlington,[18] these probabilities are then pooled to give an estimate of the overall probability that the experimentals will have a higher IQ than the controls at a given number of years after the preschool.

The second stage in the analysis will be to determine if the family variables selected for this paper are in fact correlated with IQ scores for these samples. This analysis is done by computing Pearson correlation coefficients for each project. The results will indicate whether there are significant correlations between these variables and IQ for these low SES samples or whether such correlations occur only over wider ranges of social class. The analysis will also attempt to discover whether there are any differences in correlations caused by age, sex, or experimental vs. control status.

The third and final stage in the analysis is to investigate the joint effects of preschool and the family variables by entering them all into multiple regression equations predicting IQ. Again these regressions will be constructed for each project, for each testing period. Pooled significance levels for the effects of preschool will be investigated for each of the two groups of projects established in the first stage.

In sum, the logic of the analysis runs as follows: did the projects find significant IQ differences between experimental and control children? Are the family variables correlated with IQ for these samples (i.e., are

they capable of affecting the results of preschools in some way) and are there any age, sex, or preschool effects on the correlations which should be taken into account? Finally, are the effects of preschool intervention still significant after the effect of family variables has been taken into account in a multiple regression analysis?

ANALYSIS OF THE DATA

A. Comparison of the Projects

Analysis of variance reveals that there are significant differences among the projects on all three of the variables investigated--Binet pretest, mother education, and SES. However, as seen in Table 1, the comparisons of projects using the Scheffe test show that in general, one or two projects are significantly different from most of the others but that there are no significant differences among the majority of the projects.

As shown in Table 1, the Perry Preschool and Curriculum Demonstration projects have mean pretest IQ's which are well below the other projects' (IQ's of 79.02 and 80.56 respectively, as compared with the next lowest mean IQ of 88.39). The Scheffe test reveals the Perry project to be significantly lower than every other project except the Curriculum Demonstration (CD), while the CD is significantly lower than the Palmer, Miller and Karnes longitudinal projects. The fact that the Perry and CD projects are so far below the others is due to their selection criterion which stipulated a maximum IQ of 85. These two projects must be regarded as dealing with a population of lower initial ability than the others. The other projects could be considered relatively homogeneous with repsect to the pretest.

This analysis must be treated with great caution because only the ETP, Karnes, Perry and CD projects administered genuine Binet pretests to their entire samples. Levenstein administered a Binet pretest to only about 20% of her sample (others were given the Cattel) and Palmer administered a pretest to only 60% of his sample (those who began preschool at age two were not pretested.) Miller did not administer "pretests" to her experimentals until they had already been in the programs for two to three months. Given the evidence that preschool significantly increases children's IQ's within three months,[19] therefore, it is

Table 1.

Mean Binet Pretest IQ, Mother Education and SES by Project (A)

	Binet Pretest IQ	N	Mother Education (Grade Completed)	N	Hollingshead (B) SES	N
Gordon	-----	0	10.49	115	63.67	103
ETP	88.39*	38	9.13**	88	69.68**	82
FOHV	-----	0	10.18	49	66.41*	46
Karnes Long	94.75**	102	10.13*	97	65.29*	93
Karnes P-P	89.17*	94	10.05*	91	64.68*	88
Levenstein (C)	91.18*	45	10.48	213	64.72*	212
Miller	92.37**	247	10.82	227	63.15	170
Palmer (D)	91.28**	183	11.23	299	58.17	295
W-Perry	79.02	123	9.42**	123	68.48*	122
W-Cur Dem	80.56	41	10.34	41	64.34	41
W-Carn Inf	-----	0	10.71	56	-----	0
average	89.117	937	10.46	1399	63.66	1257
overall F	12.08		10.54		12.80	
significance	.01		.01		.01	
	* significantly greater than Perry		* significantly lower than Palmer		* significantly lower SES than Palmer	
	** significantly greater than Perry and Curriculum Demonstration		** significantly lower than Palmer and Miller		** significantly lower SES than Palmer and Miller	

Notes:

(A) Significant differences between projects determined by Scheffe exact test.

(B) Since Hollingshead ranges from an "upper" SES of 11 to a lowest SES of "77", a score of 69 represents a "lower" SES than a score of 57.

(C) Only a small percent of total sample given the Binet pretest.

(D) Slightly over half of total sample given the Binet pretest.

likely that the programs had already increased the children's IQ's by the time they were "pretested." Their true pretest scores would probably have been several points lower, and the "pretest" scores reported cannot validly be compared either with control pretests or with pretests for experimentals from other projects. Furthermore, since neither Gordon, FOHV, nor the Carnegie projects gave Binet pretests (as their children entered before age three) these projects could not be included in this analysis. Thus, the analysis of the differences in pretest IQ scores applies only to a limited number of projects.

The analysis of variance on Hollingshead SES also reveals a significant difference among projects. The Scheffe test indicates that the Palmer project has a higher mean SES than the other projects.* Palmer is significantly higher than both Karnes projects, Levenstein, FOHV, Perry, and the Early Training Project. The difference is caused by the fact that the Palmer project specifically included a limited number of middle-class Black children in its design in order to achieve a greater SES range. The only other significant difference is that the Miller project is significantly higher than the ETP. This means that Gordon, CD, both Karnes projects, Levenstein, FOHV, and Perry are all "homogeneous" with respect to SES, while Miller and the ETP differ only from each other (and the ETP from Palmer).

The analysis of variance for mother education again shows a significant difference. In this case, the Perry and ETP are on the lower end of the scale (9.42 and 9.13 years of education respectively) while Miller and Palmer are on the upper end (10.28 and 11.23 years respectively). Perry and ETP are significantly lower than Miller and Palmer, while both Karnes projects are also significantly lower than Palmer. Thus, all projects except Perry, ETP,

---

* Since the Hollingshead scale uses a score of 77 to represent the lowest SES and a score of 11 for the highest SES class, a score of 53 is a higher SES than a score of 69.

and Palmer fall between 10 and 11 average years of schooling.*

In sum, the analyses reveal very little difference in the three measures among projects, but reveal for each measure two or three projects significantly different from the rest. For all projects but Perry and CD, the mean Binet pretest IQ fell between 88.3 and 94.8. For all projects but Palmer, the mean SES ranged from 63.1 to 69.7 on a scale of 11 to 77. Years of mother's education, excluding Perry, ETP, and Palmer range from 10.05 to 10.82.

Thus, the Perry and CD projects deal with children of lower IQ; Palmer deals with children of higher mean SES and mother's education; and Perry and ETP deal with children whose mothers have less education than the others. Analysis of the sample selection criteria (contained in the Appendix) confirms that the biases in the Perry, CD, and Palmer data noted above are the direct result of their selection criteria with respect to IQ or SES. An additional selection criterion is the specification by the Karnes longitudinal project of a minimum IQ of 80, which explains the fact that this project has the highest mean pretest IQ.

Procedures for assigning children to experimental or control groups varied widely, but analyses of pretest Binet, SES, and mother's education between these groups for each project reveals only one significant difference in IQ - between Palmer's control children and the group which received education at three. This difference may be due to the fact that the controls took the test at two years and eight months and the experimentals took it at three years. Given the uncertainities of the standardization of the Binet across ages and the extremely young age of the children, the difference may be, at least in part, an artifact of the test.

The analysis of attrition shows few biases between the original and final samples of any one project. However, since projects stopped testing at

* As noted previously, comparing years of education across different environments can be misleading. Therefore, a second analysis of variance was performed in which mother's education was expressed as a percentage of the median grades completed for the S.M.S.A. or city from which the children came. This analysis did not yield any substantially different results from the previous analysis.

different ages, there are biases if one compares later years across projects. An important point to note is that the Levenstein project has an artificial sort of attrition caused by the fact that many children were administered the WISC instead of the Binet in later years and hence are treated as dropouts in this analysis, even though they were tested in these years.

B. Comparison of Mean IQ Differences for Experimentals and Controls

A relevant question which should be asked before beginning the complex analysis of the effects of preschool in the context of family variables is whether a simple comparison of the average IQ for experimentals and controls indicates that preschool education affects IQ and, if so, for how long. Table 2 gives the mean IQ's and the t-test of significance for the projects whose intervention was finished by age four, while Table 3 presents the same information for projects whose intervention either began at four or continued past age four. These tables simply compare the IQ scores for the two groups at each testing period; they do not account for the effects of any initial differences in ability.

As shown in Table 2, experimental groups had higher IQ's than controls at all periods tested, although the differences were often not significant. Gordon's project shows significant differences from age four through age six (the last data collection period for the Binet). Levenstein shows significant differences at ages three and four. Her reports reveal that these significant differences continued for several years on the WISC and Binet (although too few controls were given Binets past age four to continue the analysis of Binets alone past that point). Palmer shows significant differences through age five, his last Binet testing period.

In table 3, both the Karnes and the CD projects show gains in IQ which last through post third and post second grades respectively; however, since there are no controls in these projects, their data cannot be entered

Table 2.

Mean IQ's for Experimental and Control by Project:
Gordon, FOHV, VIP, Palmer, Carnegie

		3 years	4 years	5 years	6 years
Gordon/	control	91.40	88.47	88.94	88.69
	experimental	94.88	93.89	94.02	94.01
	t	1.82	2.31	2.38	2.72
	significance	.072	.022	.018	.007
FOHV/	control	81.14	83.88	89.25	
	experimental	87.31	91.11	91.03	
	t	1.42	1.77	.36	
	significance	.162	.083	.724	
Leven-stein	control	92.72	96.0		
	experimental	101.14	105.44		
	t	3.74	3.94		
	significance	.0001	.0001		
Palmer/	control	85.32	93.31	91.13	
	experimental	91.63	97.74	96.41	
	t	3.05	2.17	2.55	
	significance	.003	.031	.011	
Car-negie	control	100.44			109.47
	experimental	104.10			111.95
	t	.82			.52
	significance	.417			.606

ıble 3.

Mean IQ's for Experimental and Control by Project

	Pretest	Post Pre Kg	Post Kg	Post 1st	Post 2nd	Post 3rd	Post 4th
ırnes Longitudinal experimental	94.75	104.88	104.52	103.38	101.38	102.98	
ırriculum Demonstration experimental	80.56	97.7	96.08	96.95	93.58		
P control	87.34	86.32	81.89	89.56	86.10		81.17
experimental	89.43	96.05	94.98	97.71	93.63		88.42
t	.74	3.96	5.30	3.07	2.47		2.63
significance	.462	.0001	.0001	.003	.016		.010
ller control	89.21	90.00	93.74	92.48	94.00		
experimental	92.87 *	97.33	94.06	92.48	89.72		
t	1.57	3.27	.14	0	-1.38		
significance	.117	.001	.891	-.999	- .176		
rry control	78.54	83.20	86.34	87.07	86.85	86.77	84.61
experimental	79.57	94.05	91.25	91.72	88.11	87.71	84.98
t	.89	5.40	2.43	2.32	.57	.43	.17
significance	.378	.0001	.017	.022	.569	.666	.861

* Not a true pretest. See text.

into this analysis. The ETP shows significant differences through fourth grade. Miller finds a significant difference only at the immediate posttest, even though she followed the children through second grade. The Perry project shows significant differences through post-first grade. As shown by Table 3, both the ETP and the Perry project experimentals are slightly, although not significantly, above the controls at the pretest. As discussed previously, the "pretest" for the Miller experimentals is not a true pretest and hence should not be compared with the pretest for the controls.

Table 4 presents the pooled significance levels for the t-tests using the technique recommended by Darlington for pooling significance levels from independent projects.[20] The significance levels were grouped by number of years after the termination of the program rather than by age for two reasons: first, if one pooled by age, it would entail pooling the results of immediate posttests with those of later followup tests (which would lead to inaccurate inference due to the "fadeout" effect often observed in preschool programs for low income children) and second, the technique of pooling only significance levels circumvents the problem of comparing IQ tests across ages. Furthermore, analyzing by the number of years after program termination answers the question of how long the effects last as opposed to the question of to what age the affects last. This table indicates that when the information from all projects is combined, there are significant differences in IQ between experimentals and controls for at least three years after the terminations of the programs. This analysis does not control for differences in pretest IQ (since for many projects this data was unavailable) and hence must be interpreted cautiously. However, if one accepts the sample selection procedures (see Appendix) as being reasonably effective in equating the experimentals and controls on initial ability, the data suggest that preschool has a significant effect on IQ which can last for several years.

Table 4.

Pooled Z's and Significance for t-tests by # Years after Preschool Termination

	significance level at immediate posttest	significance level at 1 year after preschool	significance level at 2 years after preschool	significance level at 3 or 4 years after preschool
Gordon *	.072	.022	.018	.007
ETP *	.0001	.0003	.016	.010
FOHV *	.162	.083	.724	
Levenstein	.0001			
Miller	.001	.891	-.999	-.176
Palmer *	.031	.011		
W-P	.0001	.017	.022	.569
Carnegie		.417		.606
pooled Z **	8.795	4.917	3.356	2.252
pooled significance***	.0001	.0001	.0008	.0244

* Significance level is computed from data which includes groups which had been out of the program longer than indicated, e.g., "immediate posttest" for starred projects includes some groups for whom the test was actually a one-year posttest. This procedure was chosen for two reasons: (1) it enables the reader to relate the significance levels in this table to the t-tests in tables 2 and 3 which were grouped by age rather than by number of years after the preschool, and (2) it is an extremely conservative procedure.

** $Z = \frac{\Sigma z_i}{\sqrt{k}}$ where z_i = z score for significance level from the ith project's t-test

k = number of significance tests

*** pooled significance = significance level of the pooled Z score, using a standard Z table

C. Correlation Analysis of the Relationship of Family Variables to Binet IQ

The four family variables used in this paper - mother education, SES, number of siblings and birth order - are significantly intercorrelated for most projects as shown in Table 5. It is therefore impossible to totally separate out their effects on IQ. Nevertheless, an investigation of the simple Pearson correlations of each of these variables with Binet IQ proves quite valuable in determining whether each variable is actually related to IQ for these samples. The relationship of each variable to IQ is analyzed separately in the following manner.

A separate correlation coefficient is computed for each testing period for each project. A pooled correlation coefficient is computed for each testing period from all projects which have data at that point. The basic analyses are performed with the coefficients computed for individual projects; the pooled coefficient is used only to substantiate or call into question the results obtained from a detailed analysis of the individual project coeffecients. The analysis attempts to answer four questions about the relationship between each variable and IQ:

- Does the information from all projects and from all ages point to a genuine correlation and if so, of about what magnitude?
- Is the correlation related to the age of the child?
- Is the correlation related to the sex of the child?
- Is the correlation related to whether the child attended an experimental preschool?

The first and second questions are answered primarily by analysis of graphs of of the individual project correlation coefficients, while the third and fourth are answered by t-tests for significantly different correlations for boys versus girls and for controls versus experimentals. Conclusions drawn from the

Table 5.

Intercorrelations of Family Variables by Project

	Mother education with SES	Mother education with # siblings	Mother education with birth order	SES with # siblings	SES with birth order	# siblings with birth order
Gordon	-.4811**	-.0523	-.3190*	.0789	.0252	.5868**
ETP	-.5387**	-.2026*	-.2848**	.0040	.0938	.8739**
FOHV	-.5305**	-.1865	-.2193	.2718*	.3730**	.9054**
Karnes Long	-.4847**	-.3320**	-.2615**	.3067**	.2578**	.8563**
Karnes P-P	-.5458**	-.3866**	-.3621**	.3103**	.2361*	.9317**
Levenstein	-.5382	-.2911**	-.2262**	.2906**	.2367**	.8834**
Miller	-.5075**	-.1547**	-.1193*	.2265**	.1747*	.9112**
Palmer	-.4738**	-.3872**	-.3880*	.2629**	.2715**	.9507**
W-Perry	-.2950**	-.1935*	------	-.0323	-----	-----
W-Cur Dem	.1611	.1333	.0693	.0455	-.0855	.9411**
W-Carn Inf	-----	-.1645	-----	-----	-----	-----

* $p < .05$

** $p < .01$

analysis of each variable are stated in non-technical terms prior to the actual analysis.

1. The Correlation of Mother Education with Binet IQ

Analysis of the correlation of mother education with Binet IQ for all projects leads to several observations:

- Mother's education is consistently positively correlated with IQ, although not very strongly; the evidence points to a correlation coefficient between .12 and .28.
- There is no evidence that this relation changes significantly with the child's age between the ages of three and eight.
- Only one project shows evidence of a difference between boys and girls in the magnitude or direction of the correlation.
- There is noevidence of a difference between experimentals and controls in the magnitude or direction of the correlation.

The conclusion that mother's education is positively related to child IQ should come as little surprise. What is important is that the relationship appears even in these low income populations where the range of mother's education is quite limited (only about 5% have more than a twelfth grade education). The magnitude of the correlation seems to be less than for samples with a wider range of social class, which is a result of the restricted range of mother's education in the samples.

Table 6 presents the correlations between mother education and Binet IQ for each age for each project. A glance at this table shows a confusingly wide range of values which could cause the discriminating reader to doubt the author's sanity or at least his judgment in coming to the orderly conclusions outlined above. Figure 1 graphs each of these ooefficients against the sample

size from which it was obtained.* This structuring of the information reveals that the smaller sample sizes yield wide variations in the values of the coefficients, while the correlations for larger sample sizes generally fall between .1 and .3. More importantly perhaps the correlations for the smaller sample sizes fall both above and below this range (although more fall below than above), indicating that the wider variation among the smaller samples may be due primarily to random sampling error. Specifically, the following observations can be made:

- There are no negative correlations for sample sizes greater than 50.
- There is no correlation less than 0.1 for sample sizes greater than 100.

Furthermore, inspection of Table 6 reveals that most of the negative correlations occurred in the CD project. Breaking down the correlations for this project by sex, it is found that the source of the overall negative correlations is the group of 19 females; the 22 males consistently have a positive correlation. Only this subsample of 19 girls consistently breaks the pattern of positive correlation between mother's education and IQ. The evidence for a positive correlation within these samples is further strengthened by the fact that there were 21 significant positive correlations (and only one significant negative correlation) out of 51 coefficients reported by the projects.

Figure 2 graphs the correlation coefficients by age. The solid line represents the pooled correlation coefficients for ages three through eight, while the broken lines graph the coefficients for the individual projects. As can be seen, there is

---

* Since all coefficients are included in this graph, children who were tested five times are represented on the graph five times. This procedure, in effect, biases the graph in favor of projects which tested a greater number of times. The procedure is justified for two reasons: (1) The variation in the correlations in the same project (See Table 5.) was sometimes so great that there was no legitimate method to summarize this information into a single correlation coefficient, and (2) In the author's opinion, a project which has found a consistent relationship should be weighted more strongly than one which has not found a consistent relationship.

Table 6.

Correlation of Mother Education with Binet IQ

by Project by Age

Project	Age: / Grade:	2	3	4	5	6 post KG	7 post 1st	8 post 2nd	9 post 3rd	10 post 4th
Gordon			.1267 (65)	.1189 (64)	.1152 (67)	.1189 (62)				
ITP				.0826 (83)	.1178 (86)	.0907 (84)	.1003 (79)	.1221 (77)		.1551 (77)
PCHV			-.0302 (40)	.1297 (48)	.1961 (45)					
Karnes Long.				.1607 (97)	.0737 (97)	.0534 (87)	.2453* (84)	.1457 (81)	.3741** (55)	
Karnes Pre-Post				.1100 (91)	.1972* (89)					
Levenstein		-.2247 (27)	.2289** (152)	.2532** (150)	.2506** (103)	.1047 (81)				
Miller				.1310* (226)	.1305* (223)	.2005** (213)	.1467* (203)	.1644* (186)		
Palmer			.1867** (281)	.2174** (278)	.2626** (253)					
Perry			.1816* (95)	.2358** (123)	.1136 (121)	.2683** (120)	.2141** (119)	.2937** (117)	.2061* (117)	.3404** (114)
Curriculum Demon.			-.0518 (41)	-.0311 (40)	-.1173 (41)	-.2581 (36)	-.1781 (41)	-.1924 (40)		
Carnegie			.0287 (38)			-.3276* (36)	.0514 (53)			
Pooled Correlation		-.2247 (27)	.2058** (727)	.2149** (1200)	.1793** (1125)	.1586** (719)	.1458** (587)	.1628** (502)	.2610** (172)	.2361** (191)

* $p < .05$ ** $p < .01$ Sample sizes in parentheses

CITADEL No 510
CROSS SECTION - 10 SQUARES TO INCH

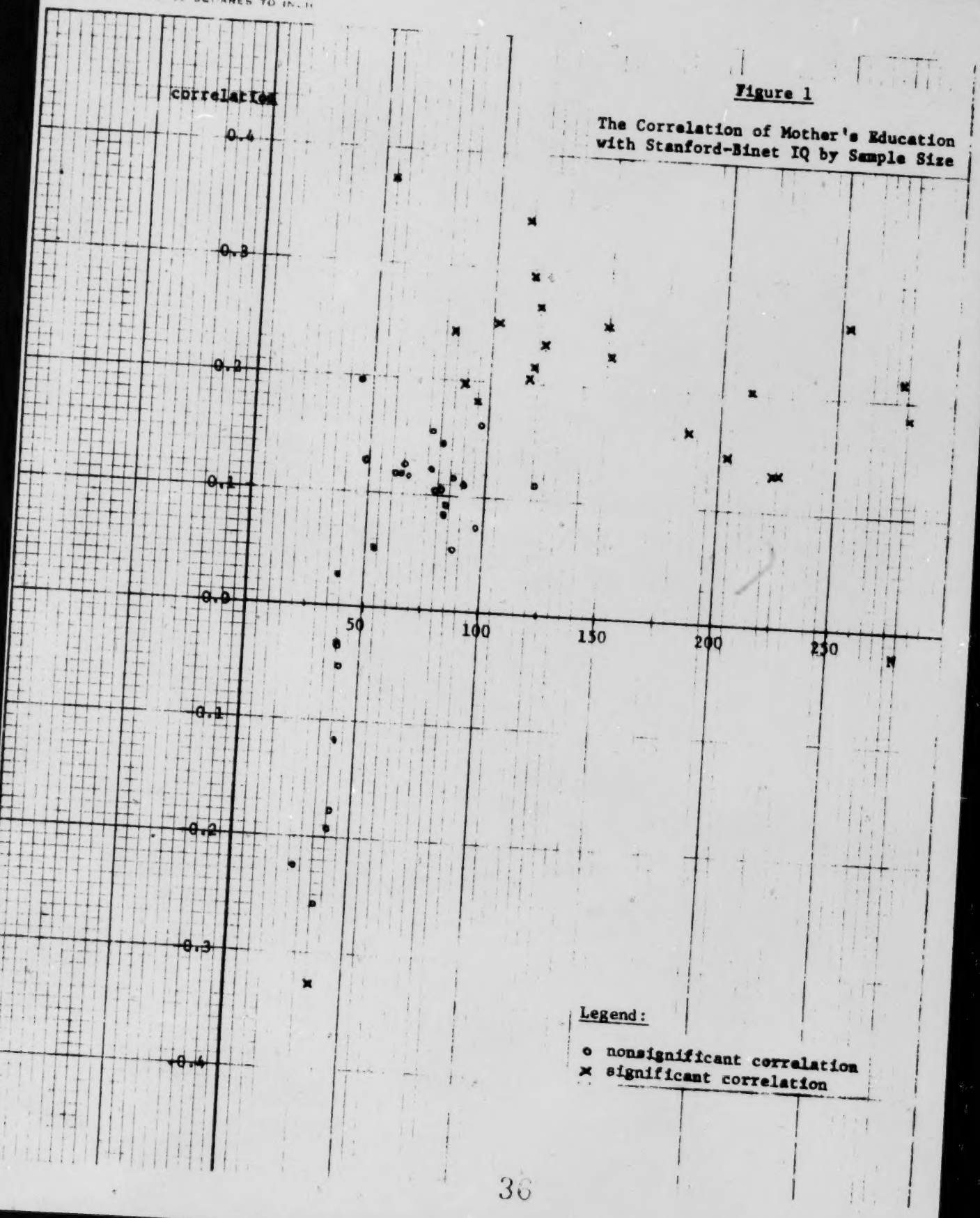
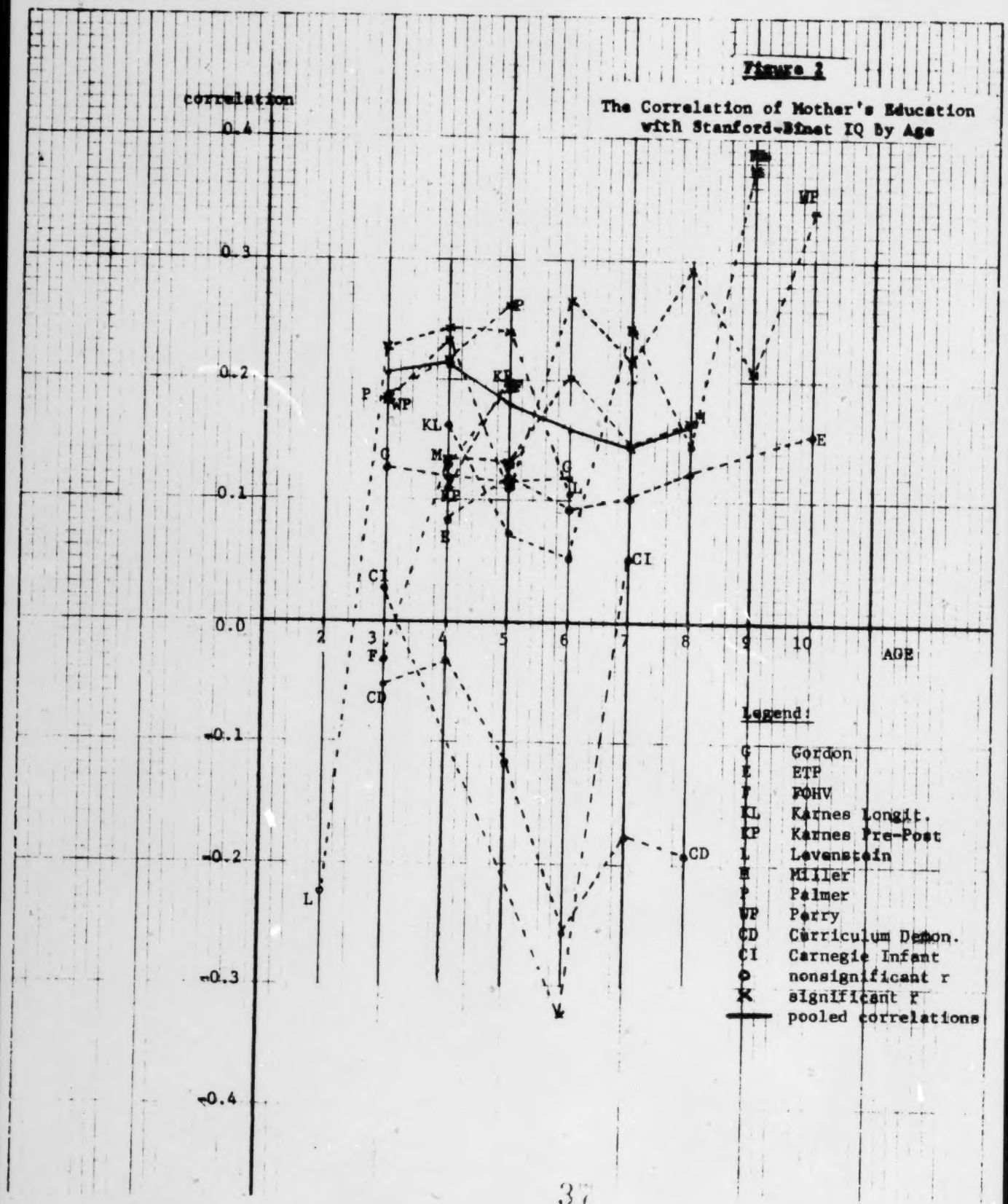


Figure 1

The Correlation of Mother's Education with Stanford-Binet IQ by Sample Size

Figure 1

The Correlation of Mother's Education with Stanford-Binet IQ by Age



no consistent trend over age in the correlations from the individual projects: some increase with age, some decrease, and some remain relatively constant. The solid line representing the pooled correlations shows a slight downward trend with age (i.e., the correlation is stronger for younger children); however, the trend is so slight that it can not be regarded as reliable given the fluctuations in total sample size over age. Thus, there appears to be no real trend over age in the correlation.

It is important to note, however, that the range of values of the pooled correlations falls within the range predicted using Figure 1 (.12 - .28). Thus, both parallel and pooled analyses point to approximately the same value for the overall correlation between mother's education and IQ - about .20.

To check for a sex difference in the relationship between the two variables, the correlations were computed separately for boys and girls, and t-tests were performed on the 48 pairs of correlation coefficients.* Only two tests were significant. At age 5, the girls' correlation of .4368 in the Karnes' pre-post project was significantly higher than the boys' correlation of -.0012 ($t=2.11$). In the CD project at age 6, the girls' correlation of -.4411 was significantly less than the boys' correlation of .0938. This project had two other pairs of correlations at different ages which revealed the same trend but were not significant. For the overwhelming majority of pairs of correlations, however, there was no difference between boys and girls.

To investigate whether attendance of an experimental preschool affects the relationship between mother's education and IQ, correlations were computed separately for experimentals and controls, and t-tests were performed on the resultant 33 pairs of correlations.** No significant differences were found. Thus, there is no

---

* Since the Palmer project had only boys, its three coefficients were excluded.

** There were no non-preschool controls in the CD project or either of the Karnes' projects. Gordon did not collect demographic information on his controls. Thus, only 33 pairs of correlations could be analyzed.

evidence that preschool in general either increases the correlation (e.g., increases IQ most for the child with a relatively well-educated mother) or decreases the correlation (e.g., increases the IQ most for the child with a less educated mother). However, this lack of evidence does not mean that the correlation is not affected by some types of preschool programs. The Gordon project, which could most reasonably be expected to affect this correlation since it employed a parent education approach, did not have the necessary data to test for a difference. The Levenstein project, in fact, which also concentrated on the mother-child interaction, shows some slight differences in the correlations (with the experimentals having a more positive correlation than the controls); however, these differences do not reach statistical significance. (The largest t-value was 1.74 while 1.96 is necessary for significance).*

In sum, then, the parallel analyses demonstrated a reasonable amount of consistency among the projects with respect to the correlation of mother's education with IQ; when taken together with the pooled correlations, these results enable one to estimate a correlation of about .2 between the two variables for lower class U.S. children. Moreover, the fact that the results of the pooled correlations are quite similar to the results of the parallel analyses indicates that the samples are quite similar. Finally, there is no evidence to suggest that the correlation varies significantly with either age (between three and eight), sex, or participation in an experimental preschool.

2. The Correlation of SES with Binet IQ

SES also appears to be related to IQ by a positive correlation** between

---

* Due to Levenstein's practive of using different IQ tests, only a part of her sample is included in these analyses; the analysis could yield different results if it were possible to include the whole sample.

** Since the Hollinshead Scale assigns a value of 11 to the highest SES and a value of 77 to the lowest, the correlations are actually negative between the scale itself and IQ, but are referred to as positive to simplify interpretation.

Table 7. Correlation of SES[1] with Binet IQ by Project by Age

Project / Age: / Grade:	2	3	4	5	6 post KG	7 post 1st	8 post 2nd	9 post 3rd	10 post 4th
Gordon		.1263 (59)	.2481* (60)	.2385* (64)	.2247* (59)				
ETP			.0549 (77)	.1586 (81)	.0856 (79)	.0596 (74)	.1195 (72)		.0536 (72)
FOHV		-.0243 (37)	.3475* (45)	.2988* (42)					
Karnes Long.			.0261 (93)	.0898 (93)	.0430 (83)	.2248* (81)	.2330* (78)	.2407* (53)	
Karnes Pre-Post			.0411 (88)	.2222* (86)					
Levenstein	.2016 (27)	.2542** (151)	.2355** (149)	.2561** (102)	.0860 (80)				
Miller			.1383* (169)	.1781* (166)	.2246** (164)	.2063** (160)	.1815* (143)		
Palmer		.1261* (277)	.1622** (274)	.2342** (249)					
Perry		-.0437 (94)	.1247 (122)	.0747 (121)	.0919 (119)	.0749 (119)	.0309 (117)	.0209 (116)	-.0387 (113)
Curriculum Demon.		.1273 (41)	.3403* (40)	.2698* (41)	.3595* (36)	.2643* (41)	.3762** (40)		
Carnegie									
Pooled Correlation	.2016 (27)	.1653** (674)	.2032** (1117)	.2071** (1045)	.1973** (620)	.1449** (483)	.1347** (451)	.1130 (169)	.0002 (185)

* $p < .05$ ** $p < .01$ Sample sizes in parentheses

1-direction of correlations reversed from that of actual Hollingshead scale.

.1 and .3 (i.e., the higher the social class, the higher the IQ). Again, there is little evidence of any effect of either age, sex, or experimental preschool on this correlation.

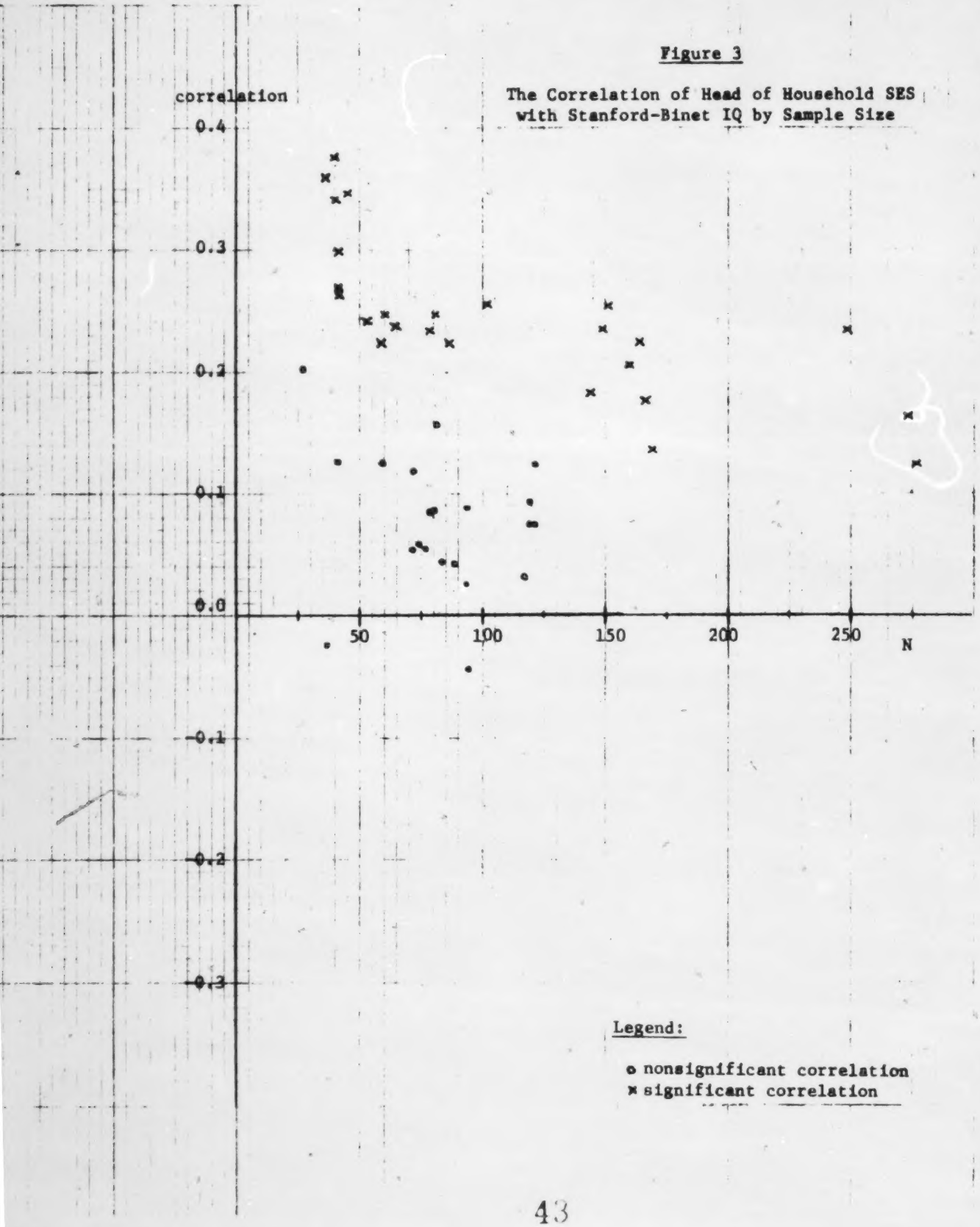
Table 7 gives the correlation coefficients at each age for each project between SES and IQ. Of 48 correlations, 25 are significant - and all of these 25 are positive. Figure 3 (the graph of correlation by sample size) reveals very strong evidence for a positive correlation between SES and IQ. Only four of the correlations reported are negative. All correlations for samples greater than 100 lie between -.04 and .26, while all correlations for samples greater than 125 lie between .12 and .26. Thus, the evidence indicates that even for these lower class samples there is still a positive correlation between social class as measured by the Hollingshead Scale, and IQ, as measured by the Stanford-Binet scale, but that this correlation is smaller than that reported across social classes. As was the case with mother's education, the smaller magnitude of the correlation is to be expected given the restricted social class range of these samples.

Figure 4 shows that there are significant correlations at every age between three and nine, and indicates no age effect on the correlation. The dashed lines representing the individual projects reveal a wide variety of trends over age, while the downward trend in the pooled correlations is not strong enough to cause one to postulate an age effect on the correlation. All pooled correlations, however, lie between .13 and .21, thus supporting the conclusions about the approximate magnitude of the correlation which were drawn from the parallel analyses in Figure 3.

The t-tests performed on the correlations of boys and girls show four significant differences out of 46 pairs of correlations. Gordon's project shows boys at age 6 to be more strongly correlated (.4559 vs. -.0701, t=2.04), while Levenstein shows the opposite trend at the same age (girls = .3230, boys = -.1779,

Figure 3

The Correlation of Head of Household SES with Stanford-Binet IQ by Sample Size

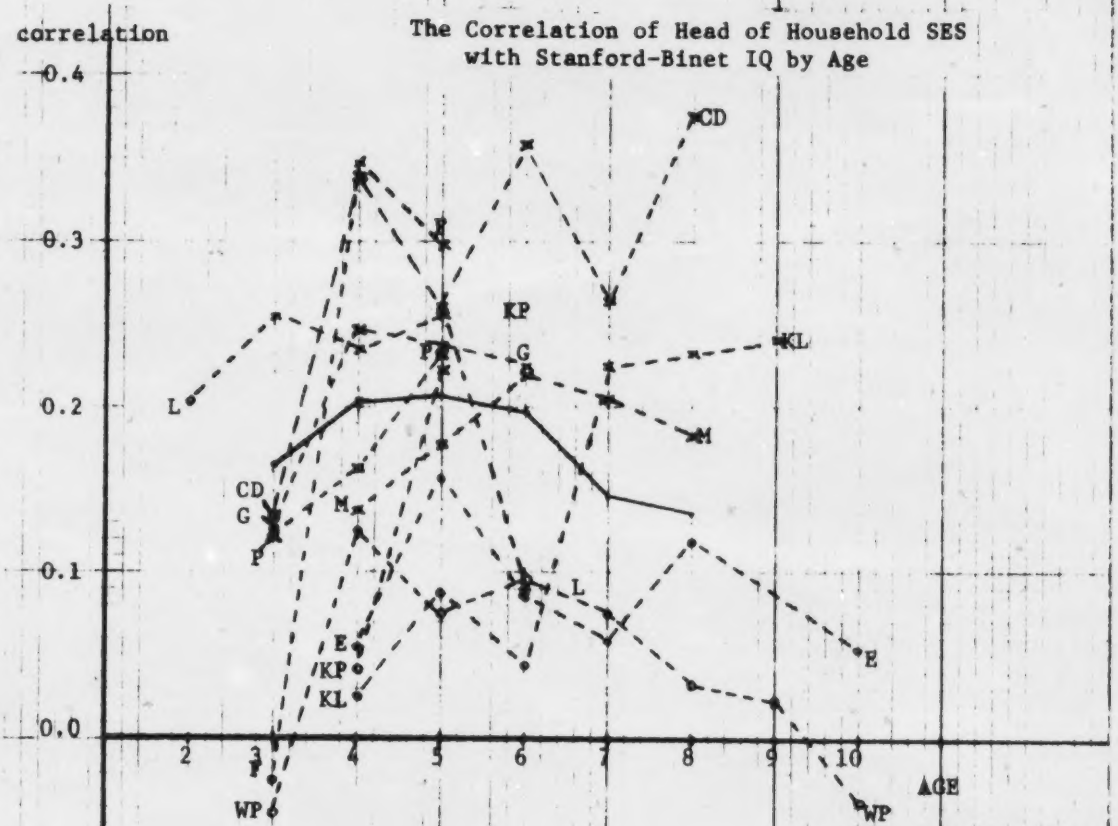


Legend:

o nonsignificant correlation
x significant correlation

Figure 4

The Correlation of Head of Household SES with Stanford-Binet IQ by Age



Legend:

G	Gordon
E	ETP
F	FOHV
KL	Karnes Longitudinal
KP	Karnes Pre-Post
L	Levenstein
M	Miller
P	Palmer
WP	Perry
CD	Curriculum Demon.
o	nonsignificant r
x	significant r
——	pooled correlations

t=2.22). The CD project also shows girls (at ages 7 and 8) to be more positively correlated (at age 7: girls = .5154, boys = -.1165, t=2.03; at age 8: girls = .5727, boys = -.0310, t=1.98). However, the overall evidence for a sex difference is not strong.

The t-tests comparing experimentals with controls show two differences in 23 pairs of correlations. The experimentals' SES scores for the Perry project at ages 8 and 10 are more negatively related to IQ than for the controls (one interpretation of this being that the program decreased the IQ disparity between high and low SES children within the sample). However, since these effects are not significant for any other project, the evidence of a general preschool effect on the correlation must be regarded as minimal.

In sum, the evidence from all projects indicates that even for lower class samples, the IQ of the children increases slightly as SES increases (correlation of about .2) and that this relationship is not significantly affected by age, sex, or participation in a preschool.

3. The Correlation of Number of Siblings with Binet IQ

Analysis of the correlations of number of siblings with IQ reveals a fairly consistent correlation of between 0 and -.2, with most of the larger ($N > 150$) samples showing correlations of about -.1. However, when separate correlations are computed for the two sexes, an interesting phenomenon occurs. Most of the projects show indications of a sex difference in the correlations; however, the projects disagree as to the direction of that difference (i.e., as to whether the boys have a stronger negative correlation than the girls or vice versa). There is slight evidence of an age effect in which the correlations grows weaker as the children grow older, but very little evidence of a preschool effect on this correlation.

Table 8 gives the correlation coefficients for each testing period for each project, and Figure 5 graphs these against sample size. This figure, even more

than Figures 1 or 3, points neatly to a single correlation coefficient (-.1) for low income populations. Not only do all large samples have correlations of about -.1, but the variability in the coefficients increases neatly in both directions (i.e., both positive and negative) as sample size decreases, indicating more random variation among the smaller samples. The figure also shows that for sample sizes greater than 100, all correlation coefficients fall between 0 and -.2, while for sample sizes greater than 150, all coefficients fall between -.04 and -.15. Thus, indications are strong that the overall correlation for lower class samples is around -.1.

In Figure 6 (correlation by age), the correlations appear to be slightly more negative at younger ages when looking across all projects, but this tendency is minimal. The pooled correlations indicate the same slight tendency. Stronger evidence, however, would be necessary before one could postulate such a trend. The values for the pooled correlations are quite close to -.1, confirming the estimate of the correlation derived from the individual project correlations.

The discovery of sex differences in the correlations of number of siblings with IQ is intriguing since the project results agree that there is a difference but disagree about its direction. Of 48 pairs of correlations, there are 11 significant differences between sexes. In 7, the girls are more negatively correlated than the boys (2 in Gordon, 5 in ETP), while in 4, boys are more negatively correlated than girls (1 in Miller, 1 in Perry, 2 in CD). Furthermore, in 10 cases a significant negative correlation for the girls is "cancelled out" by the boys' correlation so that the overall correlation (in Table 8) is not significant (1 in Gordon, 5 in ETP, 1 in Karnes' longit., 3 in Levenstein). On the other hand, in 9 cases a significant negative correlation for the boys is "cancelled out" by the girls' correlation (3 in Miller, 2 in Perry, 2 in CD). The author has no explanation for these results.

There are only two cases where the correlation differed significantly between

Table 8.

Correlation of Number of Siblings with Binet IQ

by Project by Age

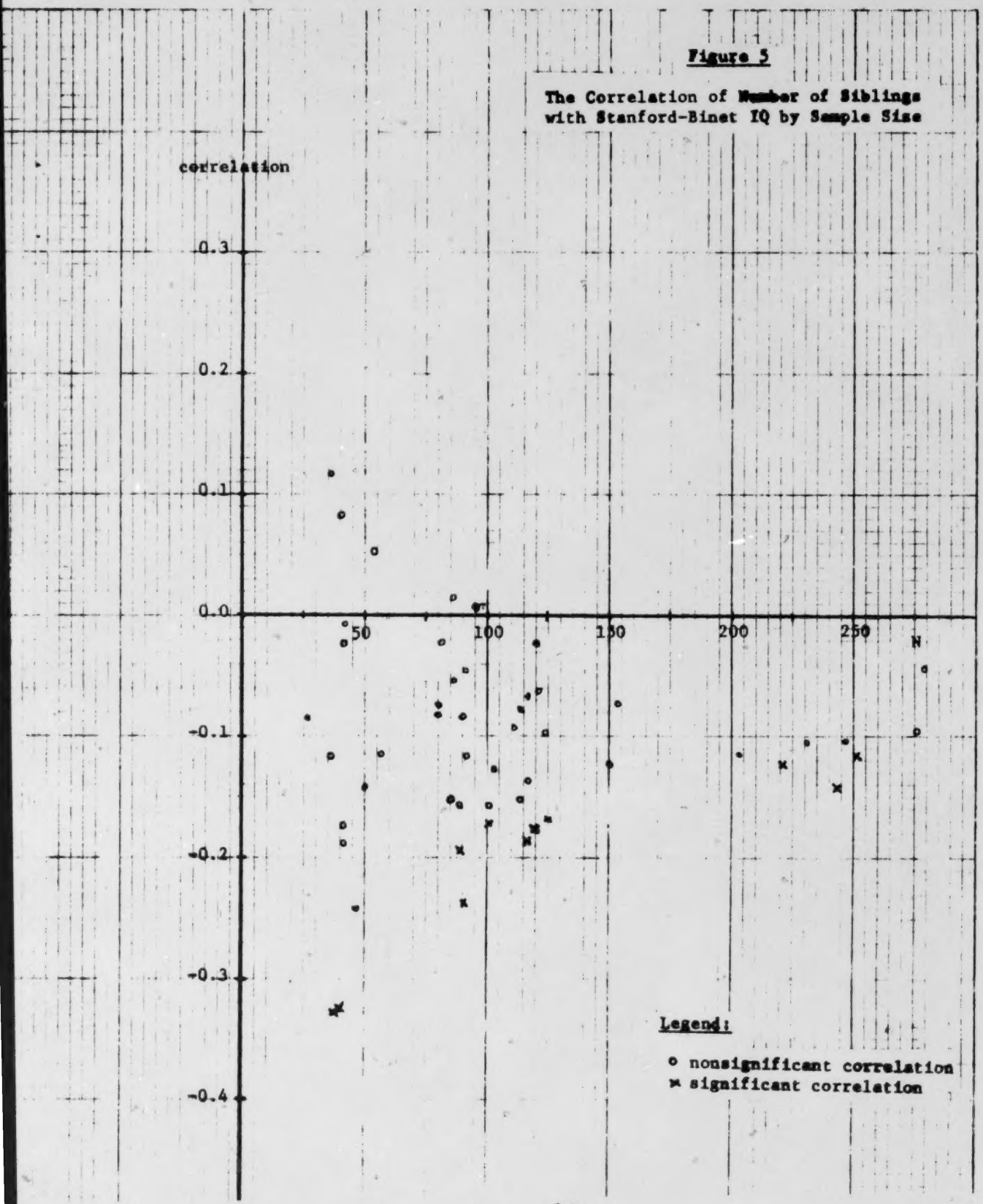
Project	Age: / Grade:	2	3	4	5	6 post KG	7 post 1st	8 post 2nd	9 post 3rd	10 post 4th
Gordon			-.0933 (111)	-.1685* (125)	-.0640 (116)	-.0778 (113)				
ETP				.0150 (86)	-.0837 (90)	-.0536 (87)	-.0226 (82)	-.0741 (80)		-.0826 (80)
FOHV			-.0072 (42)	-.1415 (50)	-.2408 (47)					
Karnes Long.				-.1721* (101)	-.1589 (101)	-.0459 (91)	-.1598 (88)	-.1517 (85)	-.1153 (57)	
Karnes Pre-Post				-.2380* (91)	-.1930* (89)					
Levenstein		-.0859 (27)	-.0725 (153)	-.1232 (150)	-.1292 (103)	-.1159 (81)				
Miller				-.1027 (247)	-.1439* (244)	-.1064 (231)	-.1228* (221)	-.1142 (203)		
Palmer			-.0411 (279)	-.0955 (276)	-.1167* (251)					
Perry			.0085 (95)	-.0980 (123)	-.0616 (121)	-.1777* (120)	-.0207 (119)	-.1380 (117)	-.1848* (117)	-.1507 (114)
Curriculum Demon.			-.1735 (41)	-.3248* (40)	-.0226 (41)	-.1158 (36)	-.1892 (41)	.0815 (40)		
Carnegie			-.3275* (38)			.1189 (36)	.0530 (53)			
Pooled Correlation		-.0859 (27)	-.1331** (774)	-.1739** (1289)	-.1228** (1203)	-.1323** (795)	-.1127** (612)	-.0947* (526)	-.1487* (174)	-.1222* (194)

* $p < .05$ ** $p < .01$ Sample sizes in parentheses

CITADEL® No. 810
CROSS SECTION 10 SQUARES TO INCH

Figure 5

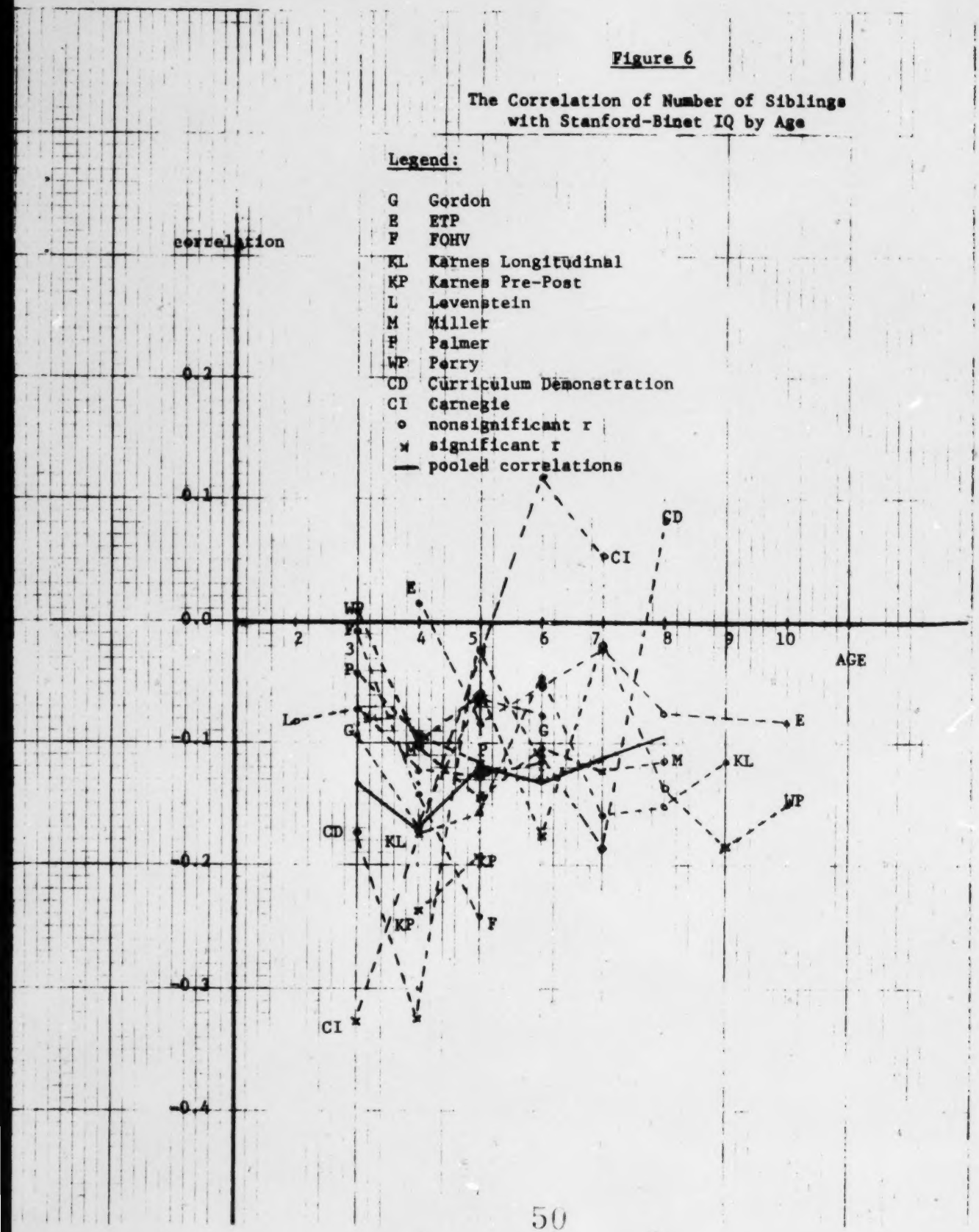
The Correlation of Number of Siblings with Stanford-Binet IQ by Sample Size



CITADEL NO. 810
... SECTION - 10 SQUARES TO INCH

Figure 6

The Correlation of Number of Siblings with Stanford-Binet IQ by Age



experimental and control children. Both are in the Perry project (at ages 3 and 4), and both showed the experimentals to be more negatively correlated than the controls (at age 3: exp. = -.2971, cont. = .2413, t=1.96; at age 4: exp. = -.2846, cont. = .0712, t=2.60). Since age three represents the pretest, these differences appear to have been initial differences and not a product of the preschool.

In sum, the overall correlation for lower class children between number of siblings and IQ appears to be about -.1. Ambiguous evidence exists about the effect of sex on this correlation, while the evidence for age and preschool effects on the correlation is minimal.

4. The Correlation between Birth Order and Binet IQ

The overall evidence for a relation between birth order and IQ is the weakest for the four variables. As might be expected given the high correlation between birth order and number of siblings, the pattern of correlation between birth order and IQ is quite similar to that between number of siblings and IQ, although the relationships in the former are weaker: probably between 0 and -.1. There are no age or preschool effects. Although there is some evidence of sex effects of the same type described for number of siblings, this effect for birth order is much weaker.

Table 10 gives the correlation coeficients, while Figure 7 graphs the correlations against sample size. Figuer 7 points toward a correlation between 0 and -.1, most probably around -.05. This correlation appears to be the weakest of the four family variables in this study.

Figure 8 presents the correlations graphed against child's age. Again, there appears to be little indication of an age effect. The values of the pooled correlations support the estimate of -.05 for the correlation.

Of 40 pairs of correlations on which t-tests were performed to discover sex effects, there were 5 significant differences: two for the Karnes' longitudinal

Table 9.

Correlation of Birth Order with Binet IQ

by Project by Age

Project	Age: / Grade:	2	3	4	5	6 post KG	7 post 1st	8 post 2nd	9 post 3rd	10 post 4th
Gordon			-.1954 (105)	-.1073 (94)	-.1040 (95)	-.1551 (91)				
ETP				.0397 (86)	-.0082 (90)	-.0281 (87)	.0489 (82)	-.0350 (80)		-.0984 (80)
FOHV			.0066 (43)	-.1444 (51)	-.1922 (48)					
Karnes Long.				.0676 (101)	-.0438 (101)	.0110 (91)	-.1458 (88)	-.1400 (85)	-.1800 (57)	
Karnes Pre-Post				-.2183* (90)	-.2295* (88)					
Levenstein		-.0595 (27)	-.0070 (153)	-.0348 (150)	-.1196 (103)	-.1319 (81)				
Miller				-.0073 (247)	-.0693 (244)	-.0294 (231)	-.0474 (221)	-.0457 (203)		
Palmer			-.0105 (279)	-.0687 (276)	-.0876 (251)					
Perry										
Curriculum Demon.			-.1714 (41)	-.2696* (40)	.0225 (41)	-.0672 (36)	-.1642 (41)	.1707 (40)		
Carnegie			-.1649 (37)			.0983 (35)	.0430 (38)			
Pooled Correlation		-.0595 (27)	-.0487 (673)	-.0847** (1135)	-.0701* (1061)	-.0350 (652)	-.0310 (478)	-.0262 (409)	-.1800 (57)	-.0984 (80)

* $p < .05$ ** $p < .01$ Sample sizes in parentheses

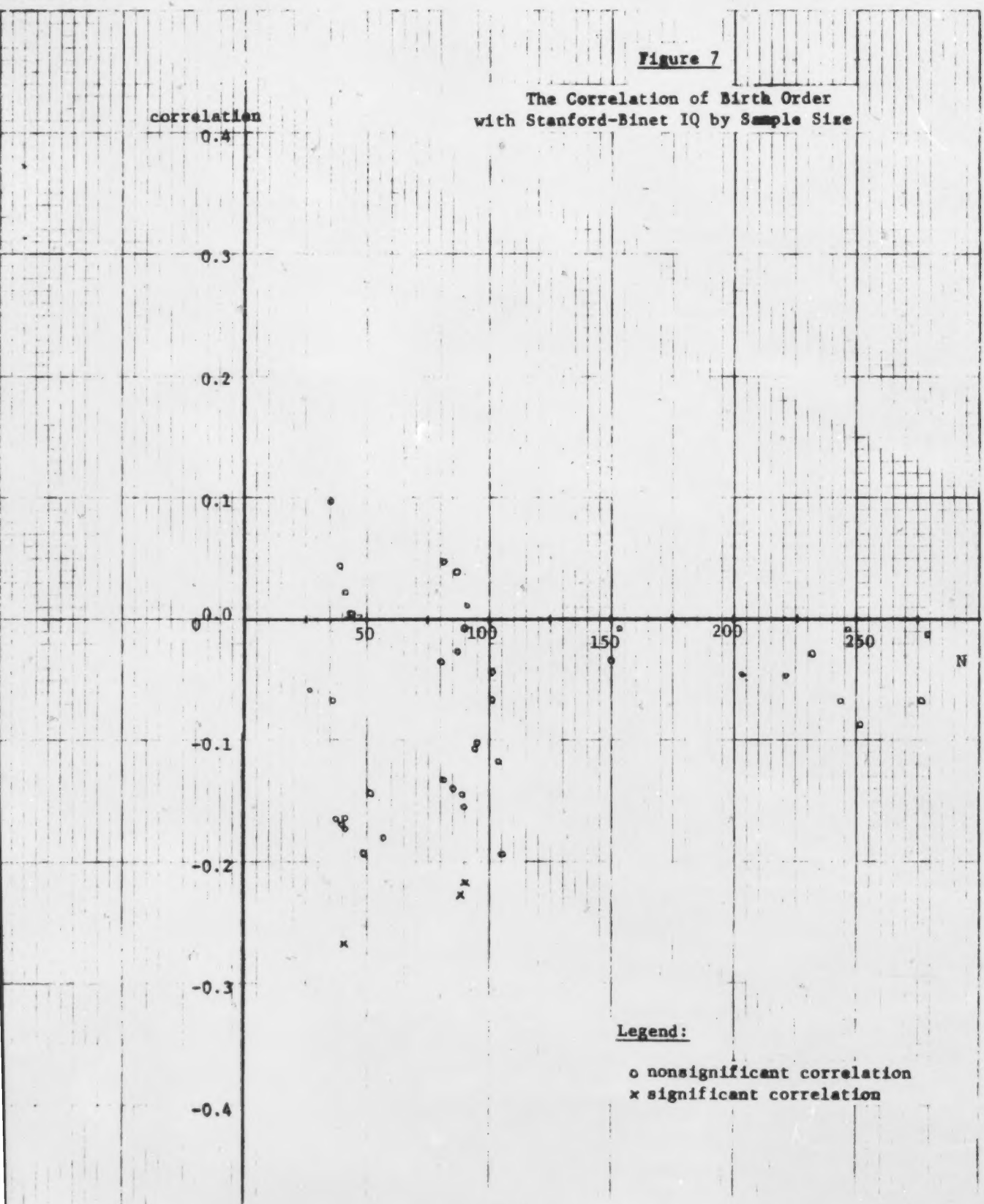
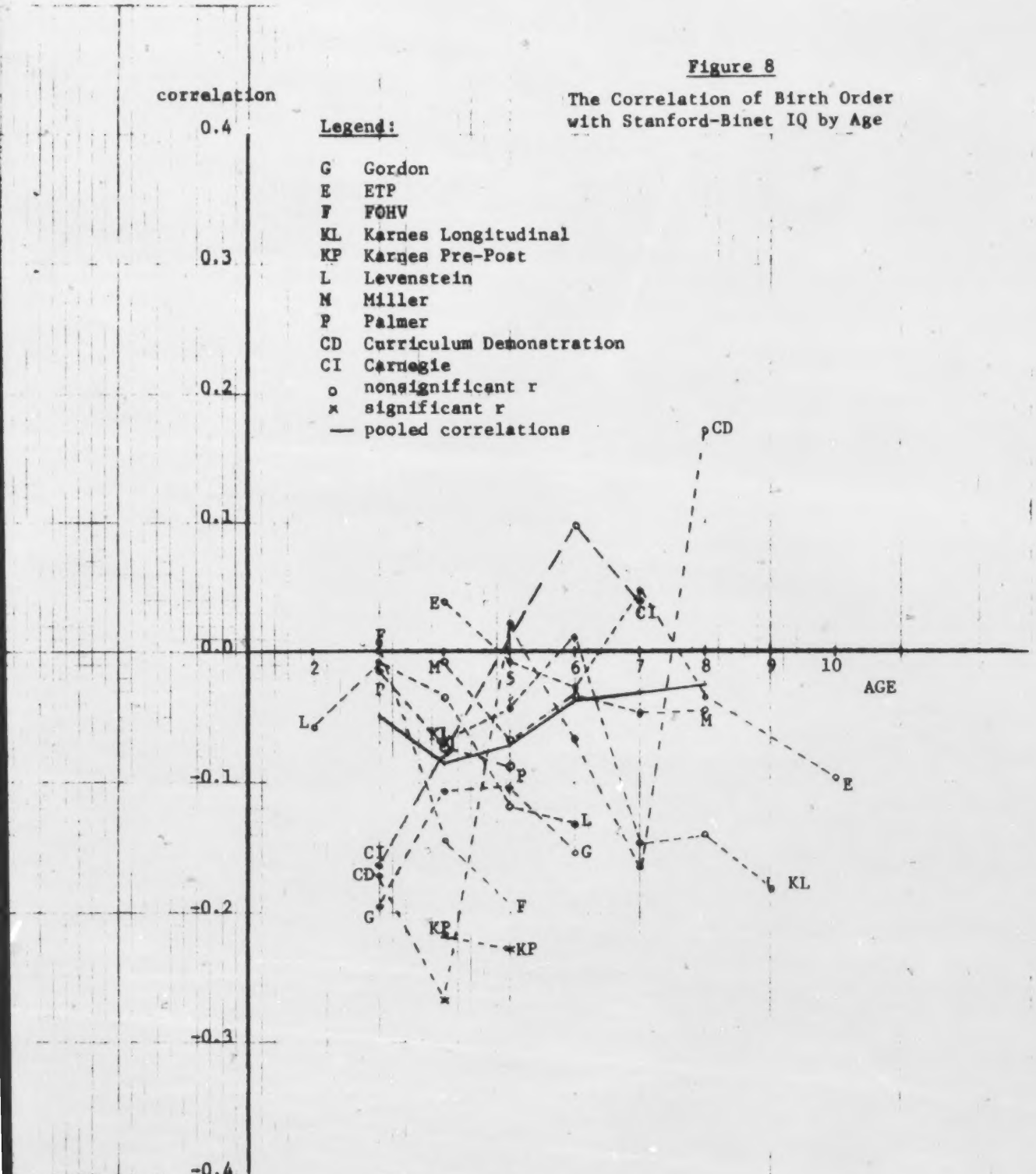


Figure 7

The Correlation of Birth Order with Stanford-Binet IQ by Sample Size

Figure 8

The Correlation of Birth Order with Stanford-Binet IQ by Age



projects (at age 4: girls =-.2484, boys = .1664, t=2.05; at age 9: girls = -.4077, boys = .1739, t=2.17), two in the CD (at age 4: girls = .3055, boys = -.6917, t=3.39; at age 6: girls = .3815, boys = -.4449, t=2.40), and one in the Miller project at age 8 (girls = .1149, boys = -.2465, t=2.56). As was the case with number of siblings, the projects disagree as to the direction of the sex effect.

Of 25 t-tests for differences between experimental and control in the correlation of birth order and IQ, there are no significant differences.

5. Conclusions from the Correlation Analysis

The correlations of the four family variables with IQ appear to be approximately as follows: .20 for mother education, .20 for SES, -.10 for number of siblings, and -.05 for birth order. These correlations are all in the same direction as those attained for samples drawn from larger ranges of social class, but they are all of smaller magnitude. This smaller degree of correlation is to be expected, since the former two variables have more restricted ranges in this sample than in the general population, while part of the correlation which the latter two show with IQ for the general population is due to their correlation to social class and this component is removed by using only lower class samples.

Thus, the contributions to explained variance of IQ by these variables for a lower class population are estimated at, respectively: .04, .04, .01, and .0025. Due to the intercorrelations, their combined contribution to the variance of IQ would probably be less than .09. Thus, although all of the variables do have a definite and measurable effect on IQ when the data from all samples is taken into account, their combined contribution to explained variance is likely to be small.*

* An observation of more peripheral interest stems from the fact that the correlations between a given demographic variable and IQ varied widely for the same children, depending on which test period was used for the correlation (i.e., there appeared to be a great deal of random fluctuation in the correlation). This result should cause one to view with some suspicion correlations published

D. Regression Analysis

The analysis to this point has shown that: (1) considering the results of all of the projects, there is a significant probability that lower class children who attend a preschool have higher IQ's than those who don't for at least three years after the preschool, and (2) the four family variables considered, although definitely correlated with IQ, appear likely to explain relatively little of the variance in IQ. The step remains of finding it, indeed, the introduction of these family characteristics into the question of whether preschool has long-term effects on IQ changes finding (1), which was derived from simple t-tests. The most appropriate vehicle for accomplishing this final step is multiple regression.

The regression formulated to ascertain the combined effects of preschool and the family variables is:

$$IQ_t = b_0 + b_1(EXP) + b_2(MED) + b_3(SES) + b_4(SIBS) + b_5(BO) + b_6(BO^2) + b_7(SEX).$$

where: IQ_t = Binet IQ at time t

EXP = dummy variable representing experimental or control status (1 = experimental, 0 = control)
MED = mother's education in grades completed
SES = Hollingshead SES scale
SIBS= number of siblings (the functional expression of family size)
BO = birth order
SEX = sex

Binet pretest scores are not included in the equation for two reasons: (1) several projects have no Binet pretest, and it would be impossible to compare equations with pretests to equations without pretests, and (2) the probable upward bias in Miller's experimental pretests and the partial use of Binet pretests by the Palmer and Levenstein projects mean that a simple covariate adjustment (i.e., entering the pretest into the regression equation) is not a universally appropriate means of control for initial ability.

Despite the high multicolinearity problems for SIBS, BO, and BO^2, all three

---

for IQ with such variables if the subjects were tested only once, especially for small N's.

are entered into the equation on the basis of the Zajonc and Markus study, which finds that a similar formulation of (family size + BO + BO^2 + d) - where "d" is a dummy variable for the last child in the family, which could not be included since our data did not alsays consist of "complete families" - is successful in explaining a significant portion of the variance of IQ.[21]

Unfortunately, the data necessary for the full equation was not collected by every project. Thus, the full equation can only be used for the ETP, FOHV, Levenstein, and Palmer projects. The Gordon project could not be used in the analysis because no demographic data was collected for the controls. The Karnes and CD projects can not be included because they have no controls. Variables left out of the equations for the other projects are as follows: Miller - SES; Perry - birth order; and Carnegie - SES. This means, in effect, that for the last three projects, the control for demographic variables is somewhat less than for the full equation.

The most complete regression equation possible is computed for each post-test period for each project: to maintain maximum comparibility, variables are not excluded from these equations unless their contribution to explained variance (R^2) is less than .001 or unless their F value for their regression coefficient is less than .01 (i.e., t was less than .1).

Since 26 equations are computed, it is not possible in this paper to completely analyze all aspects of all equations. Therefore, the analysis concentrates on two aspects of the equations - their overall explained variance (R^2) and the significance of the preschool variable within the equations.

The R^2's are generally low - ranging from a paltry .067 to a none-too-impressive .3196. The R^2 is not related to the sample size - the only R^2's of over .25 are from the ETP (N of about 80) and the Carnegie project (N of about 37), while R^2's under .10 occurr in the Miller (N of about 200), Palmer (N of about 270), Perry (N of about 120), and ETP projects. Table 10 presents the R^2 and N

for each equation. The R^2 results confirm the expectation derived from the correlation analysis that the overall contribution of these four family variables to the explained variance of IQ would be small.

The introduction of the family variables into the regression equations does not have a major impact on the significance of the preschool effect. Table 11 shows the significance levels for the regression coefficients for preschool effects compared to the significance levels for the simple t-tests of experimentals versus controls on IQ (from Tables 2 and 3). In only one case is the preschool effect significant in one procedure and not in the other: the Miller project's post-second grade testing period, in which a simple comparison of the means reveals no significant difference but in which the regression analysis controlling for family variables indicates that the experimentals are significantly lower than controls in IQ at this testing point. The other minor changes in significance levels which do occur are more prevalent in the later testing periods where the relative influence of the preschool effect when controlling for the family variables diminishes somewhat. These changes, however, are very slight.

The final question which must be answered is whether the regression analysis yields a different conclusion than that drawn from Table 4 (i.e., the conclusion that in terms of pooled significance levels, children who have participated in preschool have significantly higher IQ's for three years following termination of the intervention). Table 12 presents the analysis of pooled significance levels for the preschool regression coefficient. It is thus analogous to Table 4 with the exception that the Gordon project is excluded from Table 12. The pooled significance levels from Table 12 indicate that preschool effects last for two years, but not for three. Analysis of the individual significance levels indicates that this conclusion is not due to the effects of the family variables on the equations, but rather is due to the exclusion of the Gordon project from the regression analysis. If one makes the not unreasonable assumption

Table 10.

R^2 and Sample Size for Each Regression

Project \ Age	3	4	5	6	7	8	9	10
ETP			.266 (81)	.297 (79)	.144 (74)	.105 (72)		.098 (72)
FOHV	.184 (36)	.165 (44)	.162 (41)					
Levenstein	.205 (151)	.199 (149)						
Miller			.096 (223)	.079 (213)	.067 (203)	.090 (186)		
Palmer*	.138 (162)	.090 (272)	.135 (247)					
Perry			.214 (121)	.135 (119)	.088 (119)	.113 (117)	.094 (116)	.144 (113)
Carnegie	.320 (37)			.111 (35)	.087 (38)			

* The equation for age 3 was calculated only for the controls and for those experimentals who had been in the preschool at age 2 - not for those who began at age 3 since this would be their pretest.

Table 11.

Regression Coefficients for Effects of Preschool and Significance of Preschool Effects in the Regressions vs. Significance of Preschool Effects in t-tests.

Project	Age	Regression Coefficient for Preschool	Significance of Preschool in Regression	Significance of Preschool in t-tests on IQ
ETP	5	10.0	.0001	.0001
	6	14.26	.0001	.0001
	7	7.28	.012	.003
	8	6.36	.057	.016
	10	7.22	.024	.010
FOHV	3	7.90	.162	.162
	4	5.67	.290	.083
	5	-1.28	.800	.724
Levenstein	3	9.38	.0001	.0001
	4	8.24	.0001	.0001
Miller	5	6.33	.006	.001
	6	-1.02	-.660	.891
	7	-0.64	-.818	-.999
	8	-5.26	-.048	-.176
Palmer	3	7.68	.001	.003
	4	5.02	.015	.031
	5	6.02	.003	.011
Perry	5	10.84	.0001	.0001
	6	4.89	.013	.017
	7	4.56	.023	.022
	8	1.42	.503	.569
	9	0.80	.704	
	10	0.63	.749	
Carnegie	3	-1.06	-.80	.82
	6	0.06	.80	.52
	7	0.07	.713	.89

* The value of the regression coefficient can be roughly interpreted as the difference in IQ points between experimental and control groups when controlling for the other variables in the equation.

Table 12.

Pooled Z's and Significance for Preschool Effect in Regression by # of Years after Preschool Termination

	significance level at immediate posttest	significance level at 1 year after preschool	significance level at 2 years after preschool	significance level at 3 or 4 years after preschool
ETP*	.0001	.012	.057	.006
FOHV*	.162	.290	.800	
Levenstein	.0001			
Miller	.006	-.660	-.818	-.048
Palmer*	.0146	.0034		
Perry	.0001	.013	.023	.503
Carnegie		-.80		.80
pooled Z**	8.462	3.419	2.098	1.645
pooled significance***	.0001	.0006	.036	.5092

* Significance level is computed from data which includes groups which had been out of the program longer than indicated, e.g., "immediate posttest" for starred projects includes some groups for whom the test was actually a one-year posttest. This was done to insure comparability with Table 4.

** $Z = \frac{\Sigma z_i}{\sqrt{k}}$ where z_i = z score for significance level from the ith project's preschool regression coefficient.

k = number of significance tests.

*** pooled significance = pooled significance level of the pooled Z score, using a standard Z table.

that regression analysis would not appreciably alter the significance levels for Gordon's project from those which were obtained from the t-tests, then the pooling of significance levels from the regression results would support the finding of the t-tests that IQ effects can last for three years.

Just as important as the pooled analysis of significance levels, however, is the overall pattern of project results when examined individually. The projects are almost unanimous in concluding that preschool has a significant immediate effect on IQ. (Only the two projects with the smallest sample sizes - FOHV and Carnegie - show non-significant results at the first Binet post-test; moreover, for Carnegie, this test is one year after termination, and not an immediate post-test). The projects do, however, vary in their estimates of the longevity of effects on IQ: the ETP traces significant effects for four to five years (its last Binet data point), Gordon for three years (last Binet data point), and Weikart for two years, while Miller finds no significant IQ effects beyond the immediate post-test. Thus, while the overall indication is that statistically significant IQ effects last for several years, the projects reveal considerable variability in the persistence of IQ effects. The results point to a question which is being asked with more and more frequency: do different types of preschools have different effects on different types of children? This question will be addressed in future analyses of this data.

The pooled results of these studies, most of which are much closer to "true experiments" than the national evaluations, present a more hopeful picture with respect to long term results of preschool than do the national evaluations. Two cautions must be given in making the above statement. First, prior ability was not controlled for statistically, but was controlled for only to the extent that the sample selection procedures of the projects managed to equate the experimental and control groups. Secondly, the reader must decide for him or herself whether statistically significant IQ differences are also educationally significant.

CONCLUSIONS

In many respects the analyses presented in this paper are more a "review of the facts behind the literature" than they are genuine research. Given the problems in trying to combine the results of these studies, one should not regard the conclusions drawn from these analyses in the same light as one would regard the conclusions from a properly designed experiment of equal size. The conclusions, however, should be taken more seriously than those of reviews of the literature - if for no other reason than that the attempt to perform parallel analyses has forced the author to recognize problems of comparability which can be easily glossed over in reviews of the literature - no matter how thoroughly such reviews are performed.

Furthermore, the conclusions should be regarded as tentative because this paper represents only the first step in the larger study of what types of preschools work best for what types of children. The data base used for this paper will be expanded upon greatly both in terms of additional data from other projects and in terms of additional data from the projects already included. In particular, some data will be added on the educational experiences of the children after preschool; the addition of such data should enhance the ability to separate the effects of preschool from the effects of later education. Moreover, as the use of parallel analyses reveals both the similarities and the differences among the projects, more sophisticated analyses should be possible.

While the results of the current analysis do not by any means claim to prove that preschool has permanent effects on IQ or that a few hours a day of preschool can overcome all the debilitating effects of a child's environment, the results do indicate that, even when some of the environmental variables most commonly associated with IQ are controlled for, well run preschool programs using a great variety of approaches produce significant, immediate improvements in IQ which can last for several years. The results show that the longevity of IQ improvement

varies greatly but that the overall evidence from the projects (in terms of pooled significance levels) indicates that (assuming no significant differences in initial ability between experimental and control groups) preschool effects on IQ tend to last for at least three years. The analysis can not go beyond this length of time because three years was the last Binet data point for all but two projects.

The analysis also indicates that for lower class populations, significant correlations exist between mother education, SES, family size, and birth order and IQ but these correlations appear to be weaker than they are for the general population. Furthermore, the combined effect of these variables on the percent of variance explained for IQ is slight. Finally, there is little evidence that preschools in general affect the correlations between these variables and IQ, although there are a few instances in which such effects appear for a particular preschool project.

In sum, the analysis provides evidence that the "current wisdom" of no long term IQ effects is not universally true. Further, it indicates that the most appropriate next step in the attempt to assess the effects of preschool on the intellectual functioning of low income children is to look for the effects of different preschool approaches on different types of children.

Footnotes

1. Arthur Jensen, "How Much Can We Boost IQ and Scholastic Achievement?", Harvard Educational Review, Vol. 39, No. 1, Winter 1969, pp. 1-123.

2. Urie Bronfenbrenner, A Report on Longitudinal Evaluations of Preschool Programs: Vol. II, Is Early Intervention Effective?, DHEW (OHD) 75-25.

3. Christopher Jencks et al., Inequality: A Reassessment of the Effect of Family and Schooling in America, Harper Colophon Books, 1973, p. 86.

4. Bronfenbrenner, op. cit., and E. Kuno Beller, "Research on Teaching: Organized Programs in Early Education", in R. Travers (ed.), Handbook for Research on Teaching, Chicago: Rand-McNally & Co., 1972.

5. Ira Gordon, "What Are Effective Home Learning Environments for School Age Children?", a paper presented at the Institute for Development of Human Resources Conference, March 29, 1976.

6. See Donald T. Campbell and Albert Erlebacher, "How Regression Artifacts in Quasi-Experimental Evaluations Can Mistakenly Make Compensatory Education Look Harmful", in Jerome Hellmuth (ed.), The Disadvantaged Child, Vol. III, Brunner/Mazel, 1970, and Glen G. Cain, "Regression and Selection Models to Improve Nonexperimental Comparisons" in Evaluation and Experiment: Some Critical Issues in Assessing Social Programs, Academic Press, 1975, pp. 297-317.

7. Elizabeth Herzog et al., "Double Deprivation: The Less They Have, The Less They Learn", in Sally Ryan (ed.) A Report on Longitudinal Evaluations of Preschool Programs: Vol. I, Longitudinal Evaluations, DHEW (OHD) 75-24, 1974, pp. 69-94.

8. Lois Ellin-Data, The Effects of the Head Start Classroom Experience on Some Aspects of Child Development: A Summary of National Evaluation 1966-1969, DHEW (OHD) 76.30088, 1976.

9. R.L. Thorndike, Stanford-Binet Intelligence Scale 3rd Revision Form L-M Norms Tables, Houghton-Mifflin Co. 1973, p. 359.

10. Leon Kamin, The Science and Politics of IQ, John Wiley & Sons, 1974, p. 64.

11. Edward C. Bryant et al., Associations Between Educational Outcomes and Background Variables: A Review of Selected Literature, a monograph of the National Assessment of Educational Progress, 1974, p. 22.

12. Reported in Ibid., p. 25. Note that since these are multiple correlations, they are likely to be higher then simple correlations of a single measure of SES with IQ or achievement.

13. Martin Whiteman and Martin Deutsch, "Social Disadvantage as Related to Intellective and Language Development," in Martin Deutsch, Irwin Kety, and Arthur Jensen (eds.), Social Class, Race, and Psychological Development, Holt, Rinehart, and Winston, Inc, 1968, p. 111.

14. Reported in Arthur Jensen, "Social Class and Verbal Learning," in Deutsch et al, p. 119.

15. Jencks, op cit., p. 78

16. Bryant et al., op. cit., p. 56

17. Kenneth W. Terhune, A Review of the Actual and Expected Consequences of Family Size, National Institute of Health Publication No., (NIH) 75-779, U.S. Government Printing Office, 1974, p. 51

18. Richard B. Darlington, Radicals and Squares, Logen Hill Press, 1975, p. 525.

19. Susan W. Gray and Rupert A. Klaus, "The Early Training Project: A Seventh-Year Report", Child Development, Vol. 41, No. 4, December 1970, pp. 909-924.

20. Darlington, op. cit., p. 525

21. R.B. Zajonc and Gregory B. Markus, "Birth Order and Intellectual Development", Psychological Review, Vol. 82, No. 1, 1975.

APPENDIX

SAMPLE SELECTION AND ATTRITION

A. Sample Selection

The topic of sample selection must be divided into two subtopics: sample selection criteria and procedures for assigning subjects into experimental and control groups. The selection process for each project was quite complex, and a thorough analysis of all aspects would at least double the length of this paper. Hence, the description included here is summary in nature and, of necessity, oversimplifies the selection process of some of the projects.

The sample selection criteria used by the projects always entailed age, geographic and socioeconomic criteria of some kind. However, a number of projects employed selection criteria which distinguish them from the other projects. Palmer's project selected only boys. It also made provision for a limited number of middle class Blacks to enter the project; hence, as seen in the body of the paper, it had the highest average SES level of all projects. Weikart's Perry Preschool and Curriculum Demonstration projects employed the criterion of a maximum IQ of 85; hence, these two projects are the lowest in mean IQ, while the Karnes' longitudinal project, which utilized a minimum IQ of 80 as an entrance criterion has the highest mean pretest IQ of all the projects.

The procedures for assigning subjects to experimental and control groups differed considerably across projects and even differed within some projects, since some employed two or more types of control groups. To give an overview of the selection processes, the projects are classified under two criteria: (1) whether their controls and experimentals were selected from a common pool of subjects, and (2) whether a random or matching process was used to separate controls from experimentals. Six projects had at least one control group which was drawn from the same

pool of subjects as the experimentals: Gordon, Gray ETP, FOHV, Levenstein, Weikart Perry, Carnegie. Four projects have at least one control group which was not drawn from a common pool: ETP, Levenstein, Miller, Palmer. Of the projects where controls and experimentals were drawn from a common pool, Gordon, ETP, FOHV, and Carnegie use a randomization procedure, Perry uses matching on certain characteristics and then randomization, and Levenstein and Gordon use randomization by geographic area (Gordon used different procedures at different stages). Of the four projects which have controls not drawn from a common pool, ETP has a control in another city which was recruited solely as a control group. Levenstein uses some "after only" controls. Miller's control group consists mostly of the children on the Head Start waiting list. Palmer's controls were born one to two months later than his experimentals and were recruited specifically as controls. The CD and Karnes projects have no controls and, hence, are not dealt with here.

On the basis of the descriptions of the procedures, the only procedure which causes a certain amount of concern is that of Miller. The concern is due to the fact that the determination of who was in the control group was made by the Head Start officials who determined who was on the waiting list. Their criteria are unknown. This procedure was unavoidable given the situation of the project, and probably does not bias results since the children on the waiting list had also applied for Head Start, meaning that there is no self-selection bias involved as there might be with an "after only" control group.

Another approach to determining biases between experimental and controls is to compare their mean scores for each project on pretest IQ, SES, and mother education. These comparisons reveal that, in all projects which gave Benet pretests, the experimentals are higher than the controls at pretest, although in only one case is this difference significant - that of the Palmer project. This information,

however, cannot simply be taken at face value since (1) both Palmer and Levenstein administered a Binet pretest to only a portion of their total samples (60% and 20% respectively), and (2) the Miller pretests for experimentals appear to be biased upwards to an unknown extent because they were not administered until the children had been in the programs for two to three months. Thus, only two projects administered Binet pretests to their total sample of experimentals and controls - ETP and Perry, and in these two cases the experimentals had an advantage of 2 and 1 IQ points respectively. There are no significant differences between experimentals and controls for either SES or mother education. Moreover, there is no pattern to the differences - in some projects experimentals are better off than controls, while in other projects the reverse is true.

B. Attrition

The parameters for the analysis of attrition for each project was defined as follows:

project	criterion of dropout status	dependent variables
Gordon	existence of a Binet at age 6	Binet at age 3, SES, mother education
ETP	" " " " " " 10	" " pretest " " "
FOHV	" " " " " " 5	" " " 3 " " "
Karnes long	" " " " " " 9	" " pretest " " "
Karnes p-p	" " " " " " 5	" " " " " "
Levenstein	" " " " " " 5	" " age 3 " " "
Miller	" " " " " " 8	" " pretest " " "
Palmer	" " " " " " 5	" " age 3 " " "
Perry	" " " " " " 10	" " pretest " " "
CD	" " " " " " 8	" " " " " "
Carnegie	" " " " " " 7	" " age 3 " "

Attrition was analyzed separately for each project in three stages:

(1) X^2 analysis of differential attrition rates between experimental and control groups.

(2) t-tests comparing dropouts with the final sample on initial Binet IQ, mother education, and SES.

(3) 2X2 analyses of variance using initial Binet IQ, mother education, and SES separately as dependent variables and dropout status and experimental/control status as the independent variables.

Only two projects have significantly different rates of attrition between their experimental and control groups: Levenstein, and Palmer who had significantly greater attrition in the control groups. The attrition for Levenstein is an artificial one created by the stipulation of this analysis that a subject must have a Binet score. Much of her "attrition" is caused by the fact that many of her subjects were given WISC's rather than Binets in later follow-up periods.

Differential rates of attrition do not necessarily pose a problem of major magnitude unless it can be demonstrated that either: (1) those who dropped out were significantly different on some relevant characteristic from the final sample, or (2) the attrition pattern in the control group with respect to some relevant characteristic was significantly different from the attrition pattern in the experimental group. The first question is answered by comparing the average initial IQ's, mother education, and SES of dropouts to the final sample by means of t-tests. The second is answered by means of the two way analysis of variance described above.

The t-tests reveal no significant differences between dropouts and final samples on initial IQ for any of the projects. Two projects have significant differences in mother education: Karnes' longitudinal and Carnegie. For age 9, Karnes' final sample has a significantly lower education than the dropouts (final = 9.8, dropouts = 10.6, t = .029); however, the difference is not significant for attrition

at age 8. (This effect is due to the fact that two of Karnes' groups were not tested at age (and that, therefore, attrition is much greater at that age than at age 8). In the Carnegie project the mothers of dropouts have a significantly higher education than the mothers of the final sample; however, there were only three dropouts (one of whom had a mother who had had two years of college). The t-tests on SES revealed two projects where there was a significant difference between dropouts and the final sample. Karnes' longitudinal project showed that dropouts had a significantly higher SES rating, while Levenstein's "attrition" group is significantly lower in SES than her final Binet sample. It should be noted, however, that although these significant differences in the Hollingshead SES score do occur, the average scores for these groups are all in the same Hollingshead SES category (category five - lower class).

Overall, attrition causes very little bias within the individual projects; however, since different projects stopped testing at different times there are attrition effects across projects. Since this paper emphasizes parallel rather than pooled analyses, no overall pooled attrition analysis is attempted.

The two way analyses of variance also reveal very little attrition within projects. Of the analyses for each project on initial Binet IQ, SES, and mother education, there are no significant results. (The analysis of mother education for the Gordon project has a significance level for the overall F of .026, with a significance level for the interaction effect of propout status and preschool status of .031; however, two of the cells had only one observation each so that the analysis of variance cannot be properly interpreted).

In sum, then, the analysis of attrition reveals no serious problems which need be adjusted for or otherwise taken into account in interpreting the results of the analyses of individual projects in this paper. Tables giving full results of the attrition analyses are available from the author upon request.

Bibliography

Beller, E. Kuno, "Impact of Early Education on Disadvantaged Children", in Say Sally Ryan (ed.) A Report on Longitudinal Evaluations of Preschool Programs: Vol. I, Longitudinal Evaluations. DHEW Publication No. (OHD) 75-24, 1974, pp. 15-48.

---------- "Research on Teaching: Organized Programs in Early Education", in R. Travers (ed.) Handbook for Research on Teaching. Rand-McNally & Co., 1972.

Bryant, Edward C. et al., Associations Between Educational Outcomes and Background Variables: A Review of Selected Literature, A Monograph of the National Assessment of Educational Progress, 1974.

Bronfenbrenner, Urie, A Report on Longitudinal Evaluations of Preschool Programs: Vol. II, Is Early Intervention Effective?. DHEW Publication No. (OHD) 75-25, 1974.

Cain, Glen G., "Regression and Selection Models to Improve Nonexperimental Comparisons", in Evaluation and Experiment, Some Critical Issues in Assessing Social Programs, Academic Press, 1975.

Campbell, Donald and Erlebacher, Alfred, "How Regression Artifacts in Quasi-Experimental Evaluations Can Mistakenly Make Compensatory Education Look Harmful", in Jerome Hellmuth (ed.), The Disadvantaged Child, Vol. 3, Brunner/Mazel, Inc., 1970, pp. 185-210.

Cicirelli, Victor G., "Sibling Structure and Intellectual Ability", Developmental Psychology, Vol. 12, No. 4, 1976, pp. 369-370.

Darlington, Richard B., Radicals and Squares, Logen Hill Press, Ithaca, New York 1975.

Deutsch, Martin, Katz, Irwin, and Jensen, Arthur, Social Class, Race, and Psychological Development, Holt, Rinehart, and Winston, Inc., 1968.

Deutsch, Martin, et al., An Evaluation of the Effectiveness of An Enriched Curriculum in Overcoming the Consequences of Environmental Deprivation, DHEW, Grant No. OE-5-10-045, 1971.

Ellin-Data, Lois, The Effects of the Head Start Classroom Experience on Some Aspects of Child Development: A Summary of National Evaluations 1966-1969, DHEW Publication No. (OHD) 76.30088, 1976.

Gray, Susan and Klaus, Rupert, "The Early Training Project: A Seventh-Year Report", Child Development, Vol. 41, No. 4, Dec. 1970, pp. 909-924.

Gray, Susan, Ruttle, Kristi, and Wyrick, Michael, "Home Visiting with Mothers of Toddlers and Their Siblings: Interim Report on Second Post-Testing of Waves I & II, A Report to the National Institute of Education on Contract No. NE-C-00-3-0261", 1975.

Gordon, Ira, "What Are Effective Home Learning Environments for School Age Children?", a paper presented at the Institute for the Development of Human Resources Conference, Gainesville, Fla., March 29, 1976.

Gordon, Ira and Guinaugh, Barry, A Home Learning Center Approach to Early Stimulation, Institute for Development of Human Resources, University of Florida, 1974.

Herzog, Elizabeth, Newcomb, Carol, and Cisin, Ira, "Double Deprivation: The Less They Have, The Less They Learn", in Sally Ryan (ed.), A Report on Longitudinal Evaluations of Preschool Programs: Vol. I, Longitudinal Evaluations. DHEW Publication No. (OHD) 75-24, 1974.

Jencks, Christopher, et al., Inequality: A Reassessment of the Effect of Family and Schooling in America, Harper Colophon Books, 1973.

Jensen, Arthur, "How Much Can We Boost IQ and Scholastic Achievement", Harvard Educational Review, Vol. 39, No. 1, Winter 1969, pp. 1-123.

Kamin, Leon, The Science and Politics of IQ, John Wiley & Sons, 1974.

Karnes, Merle, Zehrbach, Reid, and Teska, James, "The Karnes' Preschool Program: Rationale, Curricula Offerings, and Follow-up Data", in Sally Ryan (ed.) A Report on Longitudinal Evaluations of Preschool Programs: Vol. I, Longitudinal Evaluations. DHEW Publication No. (OHD) 75-24, 1974.

Madden, John, Levenstein, Phyllis, and Levenstein, Sidney, "Longitudinal IQ Outcomes of the Mother-Child Home Program", Child Development, Vol. 47, No. 4, Dec. 1976, pp. 1015-1025.

Miller, Louise, and Dyer, Jean, Four Preschool Programs: Their Dimensions and Effects, a Monograph of the Society for Research in Child Development, Serial No. 162, Nos. 5-6, October 1975.

Palmer, Francis, "Final Report: The Effects of Minimal Early Intervention on Subsequent IQ Scores and Reading Achievement", a paper presented at the American Psychological Association Conference, Sept. 3, 1976.

Seitz, Victoria, "Long-Term Effects of Intervention: A Longitudinal Investigation", a paper presented at the American Psychological Association Conference, Sept. 3, 1976.

Terhune, Kenneth, A Review of the Actual and Expected Consequences of Family Size, National Institute of Health Publication No. (NIH) 75-779, 1974.

Thorndike, Robert, Stanford-Binet Intelligence Scale 3rd Revision Form L-M Norms Tables, Houghton-Mifflin Co., 1973.

Weikart, David, et al., Longitudinal Results of the Ypsilanti Perry Preschool Project, High/Scope Educational Research Foundation, Ypsilanti, Mich. 1970.

Weikart, David, et al., Home Teaching with Mothers and Infants, High/Scope Educational Research Foundation, Ypsilanti, Mich. 1974.

Zajonc, R.B., and Markus, Gregory, "Birth Order and Intellectual Development", Psychological Review, Vol. 82, No. 1, 1975, pp. 74-88.